

T.C.
YEDITEPE UNIVERSITY
GRADUATE INSTITUTE OF SOCIAL SCIENCES

**Optimization of the Cut-off Points in Bankruptcy Prediction
in Turkish Companies**

By
Cenk Toker

Advisor
Asst. Prof. Dr. Çağatay AKARÇAY

Submitted to the Graduate Institute of Social Sciences in
Partial Fulfilment of the Requirements for the Degree of
Doctor of Philosophy in Finance

ISTANBUL, 2021

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Cenk TOKER

Signature:

ABSTRACT

Prediction of bankruptcy using financial ratios is a popular area of research for academicians as well as the managers, creditors and investors. Most of the research about bankruptcy prediction aim to maximize the number of companies correctly classified, without considering the costs related with the errors of misclassification.

The purpose of this study is to minimize the expected costs of bankruptcy by considering different costs in the Type I and Type II error regions. The research question - if the optimal cut-off point considering to minimize the costs in Type I and Type II Error Regions in Bankruptcy Prediction is equal to optimal cut-off point considering to minimize the number of companies in Type I and Type II Error Regions- is answered. Non-financial companies listed in Borsa Istanbul between 1998 to 2015 are categorized according to their bankruptcy status and a logistic prediction model is constructed. After the validation of the logistic model, the effect of the cut-off point change on the expected cost of misclassification is measured.

The logistic model is proved to be successful in predicting the bankruptcy 2 years before it occurred. And it is also observed that cut-off point is important in minimizing the expected costs of bankruptcy and it is different than the cut-off point to maximize the number of correctly classified companies.

ÖZET

İflasın tahminlenmesi, akademisyenler için olduğu kadar, şirket yöneticileri, borç verenler ve yatırımcılar için de popüler bir araştırma konusudur. İflas tahminleme ile ilgili çalışmaların çoğu, yanlış sınıflandırma (misclassification) sebebiyle ortaya çıkabilecek maliyetlerin ne olduğundan bağımsız, tahminleme sonucu doğru sınıflandırılan şirket adedini maksimize etmeyi amaçlar.

Bu çalışmanın amacı, tahminleme esnasında ortaya çıkan Tip I ve Tip II hata bölgelerindeki birbirinden farklı yanlış sınıflandırma maliyetlerini göz önüne alarak, iflas tahminlemesinde beklenen yanlış sınıflandırma maliyetlerini (expected misclassification costs) minimize etmeyi amaçlamaktadır. Tip I ve Tip II hata bölgelerinde ortaya çıkacak yanlış sınıflandırma maliyetini minimize eden kesim noktası ile aynı bölgelerdeki şirket adedini minimize eden kesim noktası değerlerinin eşit olup olmayacağını araştırır. Borsa İstanbul’da 1998 – 2015 yılları arasında listelenmiş, finansal sektörler dışındaki şirketler iflas durumlarına göre sınıflandırılarak, lojistik model oluşturulmuştur. Modelin doğrulanmasından sonra, kesim noktasındaki değişikliğin, beklenen yanlış sınıflandırma maliyetleri üzerindeki etkisi ölçülmüştür.

Lojistik modelin, iflası 2 yıl önceden tahminlemede başarılı olduğu görülmüştür. Ayrıca, kullanılan kesim noktasının, iflas tahminlemede ortaya çıkabilecek beklenen maliyet üzerinde önemli etkisinin olduğu ve bu kesim noktasının, tahmin sonucunda doğru sınıflandırılan şirket sayısını maksimize etmeye yönelik bir kesim noktasından farklı olabileceği görülmüştür.

ACKNOWLEDGEMENT

This research would not have been possible without the support of

My advisor Asst. Prof. Dr. Çağatay AKARÇAY with his endless support and positive approach in the completion of this dissertation,

Members of my dissertation committee, Prof. Dr. Dursun Arıkboğa and Prof. Dr. Ahmet Ozcam with their valuable comments and feedback that made this study develop from its initial stage to the end.



CURRICULUM VITAE

Cenk Toker received his Bachelor of Science Degree in Electrical Engineering from Istanbul Technical University in 1997 and his Master of Business Administration Degree in Finance and Accounting from Yeditepe University in 2007.

He is also in active business life since 2000, obtaining several management positions related with Customer Relations and Marketing, in Telecommunication, Internet and Pay TV companies.



TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENT	v
CURRICULUM VITAE	vi
TABLE OF CONTENTS	vii
List of Figures	viii
List of Tables	ix
 1. INTRODUCTION	 1
1.1. Importance of the Study	1
1.2. Outline of the Study	3
 2. LITERATURE REVIEW	 4
2.1. The Ground Work For The Usage of Financial Ratios in Academic Literature	4
2.2. The Usage of Financial Ratios in Academic Literature	11
2.2.1. Descriptive Purposes	12
2.2.2. Decision Making Purposes & Methods	14
2.2.3. Predictive Methodologies Using Financial Ratios	7
2.2.3.1. Prediction of Failure or Bankruptcy	17
2.2.3.2. Prediction of Other Phenomenon	27
2.2.4. Predictive Methodologies Using Market Variables	40
2.3. Development of Literature about Prediction of Failure	45
2.3.1. Prediction of Failure with Cash Flow Basis Accounts	46
2.3.2. Prediction of Failure with Probabilistic Models	59
2.3.3. Hazard Model and Prediction of Failure with Time Series Data	76
2.3.4. Evaluations and Re-tests of the Known Models	77
2.3.5. Advanced Techniques in Prediction of Failure using Artificial Intelligence and Machine Learning	79
2.4. Costs Related with Bankruptcy	93
2.4.1. Direct Costs of Bankruptcy	96
2.4.2. Indirect Costs of Bankruptcy	96
2.4.2.1. Indirect Costs Calculation with Regression Analysis	98
2.4.2.2. Indirect Cost Calculation as measured by Security Analysts	100
2.4.2.3. Expected Bankruptcy Costs vs Tax Benefits Trade off	101
2.5. Legal Definitions of Bankruptcy	120
2.5.1. Turkish Legal Perspective	120

2.5.1.1.	According to the way the bankruptcy is declared	120
2.5.1.2.	According to the Laws and Codes about bankruptcy	120
2.5.1.3.	According to the initiator of the bankruptcy process	121
2.5.1.4.	According to the debt is due or not	121
2.5.2.	American Legal Perspective	121
3.	METHODOLOGY	123
3.1.	Selection of the Companies	123
3.1.1.	Bankrupt Companies Selection	124
3.1.2.	Non-bankrupt Companies Selection	128
3.2.	Independent variables	128
3.3.	Selection Bias	132
3.3.1.	Choice Based Selection Bias	132
3.3.2.	Complete Data Criterion	133
3.4.	Outlier Definitions and Data Elimination	134
3.4.1.	Outliers	134
3.4.2.	Economic Crisis Period	134
3.5.	Data Construction	135
3.6.	Statistical Method	136
4.	RESULTS AND FINDINGS	138
4.1.	Mean Values	138
4.2.	Multicollinearity Between Independent Variables	140
4.3.	Construction of the Logistic Model	143
4.3.1.	Stepwise Regression	143
4.3.2.	Classification Accuracies	146
4.4.	Application of the Koh Expected Misclassification Cost Model	151
5.	LIMITATIONS AND FURTHER RESEARCH	159
6.	CONCLUSION	161
7.	REFERENCES	163

List of Figures

Figure 2.1. Daily and Cumulative Returns Before and After Bankruptcy Filing	43
Figure 2.2. Layers in a Neural Network	81
Figure 2.3. Combination of Factors in the Neuro Fuzzy Model	84
Figure 2.4. Bank Asset Size vs Loan Loss Recovery Rates	113

Figure 3.1. Logistic (Sigmoid) Curve	137
Figure 4.1. Type I and Type II Error Regions Representation	147
Figure 4.2. The ROC Curves	150
Figure 4.3. Expected Misclassification Costs vs Different Cut-offs	155
Figure 4.4 3D View of Expected Misclassification Costs	156
Figure 4.5. Optimal Cut-offs and Minimized Expected Misclassification Costs	157

List of Tables

Table 2.1. Prediction of the Mean Values of Failed and Non-failed Firms	19
Table 2.2. Accuracies of the Different Financial State Estimations of Lau's MLA	30
Table 2.3. Dopuch et al. Ratios Only Model	37
Table 2.4. Dopuch et al. Ratios and Auditor Opinion Model	38
Table 2.5. Total Variance of Returns for Failed and Non-failed Firms	41
Table 2.6. Aziz et.al. Cash Flow Based Partial Logit Model Coefficients	54
Table 2.7. Comparative classification accuracy between Aziz et.al CFB model, Altman et.al. ZETA and Altman Z model	54
Table 2.8. Comparative Classification and Prediction Accuracy Between Aziz and Lawson Cash Flow Based Model, Mixed Model, Altman Z, and Altman et.al. Zeta Models	57
Table 2.9. t Ratios to Compare the Accuracies of Bankruptcy Prediction Models	58
Table 2.10. Prediction Results of Ohlson Model	63
Table 2.11. Factor Loadings of the Pinches et al. study	65
Table 2.12. Summary of Sample Selection / Data Collection Procedures Used in Financial Distress Research	69
Table 2.13. Comparison of the Zmijewski Unweighted and WESML Probit Model Estimations Across Alternative Estimation Samples	72
Table 2.14. Comparison of the Zmijewski Unweighted and WESML Probit Model Classifications Across Alternative Estimation Samples	74
Table 2.15. Comparison of the Zmijewski Unweighted and WESML Probit Model Classifications in the Prediction Sample Across Alternative Estimation Samples	75
Table 2.16. Summary of Factor Analysis of Chen et.al. Neuro Fuzzy Model	83
Table 2.17. Results of Chen et.al. Neuro Fuzzy Model in training sample	85
Table 2.18. Results of Chen et.al. Neuro Fuzzy Model in testing sample	85
Table 2.19. The accuracies of Chen Model with different level factors across different algorithms	86
Table 2.20. The Comparison of the Different Algorithms in Son et.al. Study	92
Table 2.21. Change of the Market Value and Bankruptcy Costs (in Railroad Companies) as Bankruptcy Approaches	95
Table 2.22. Direct and Indirect Costs of Bankruptcy vs the Company Value	100
Table 2.23. Comparison of the Tax Benefits from PV of Bankruptcy Costs vs. PV of Interest on Debt and Capitalized Leases (At t-2 Prior to Bankruptcy)	103
Table 2.24. Comparison of the Tax Benefits from PV of Bankruptcy Costs vs. PV of Interest on Debt and Capitalized Leases (At t-1 Prior to Bankruptcy)	104
Table 2.25. Comparison of the Classification Accuracies of ZETA and Z Models	106

Table 2.26. Recovery Rates on Loan Losses	113
Table 2.27. Collection and Recovery Experience on Sample of Charged-off Loans	115
Table 2.28. Results of the Koh Logit Model	118
Table 2.24. Summary of Optimal Cut-off Points in Koh Probit Model	119
Table 3.1. Reasons for Elimination from the Bankrupt Company List	127
Table 3.2. List of Bankrupt Companies	129
Table 3.3. List of Nonbankrupt Companies	130
Table 3.4. List of Independent Variables	131
Table 4.1. Mean Values of the Independent Variables	141
Table 4.2. Multicollinearity Between Independent Variables	142
Table 4.3. Signs of the Independent Variables	144
Table 4.4. Coefficients of the Final Model For Each Sample	144
Table 4.5. Classification Accuracies	149
Table 4.6. AUC Values of Different Samples	151
Table 4.7. Estimated Misclassification Costs of Different Samples	153
Table 4.8. Expected Misclassification Costs vs Different Cut-offs	158

1. INTRODUCTION

1.1 Importance of the Study

There are many areas that the results of the Bankruptcy Prediction studies are used by many different benefactors. Bankruptcy Prediction has been a favourite topic among academicians and business practitioners.

The outcomes of the research are important for the Managers because, to avoid failure, they are warned about their companies' performance indicators which they have to improve.

The outcomes are important for Creditors to evaluate the costs of their lending decisions and credit scoring and the risks of the borrower companies.

Investors use these outcomes to decide where to invest.

Auditors should definitely be using the outcomes to get a clear picture of the company audited, in their going and non-going concern opinions.

Recent failures in last decade in predicting bankruptcies of giant companies (Enron, Worldcom) caused a turmoil and mistrust in the financial markets, and the abilities of the auditors were in question. As a consequence of this mistrust, Policy Makers and Regulators use the outcomes to establish the required standards for the companies and industries.

The improvement of technology and the increase in the computational power supports the improvement of the research about Bankruptcy Prediction. Since it is a prediction problem, in recent years, the data and computer engineers are also interested in the area of research, as well as the social science researchers.

Although Bankruptcy Prediction is a popular area of research, most of the studies aim to maximize the number of companies correctly classified as bankrupt or non-bankrupt, or minimize the number of companies in Type I and Type II Error regions. There are some studies that also consider the quality of the prediction, that is to say, the costs related to the misclassification in the models. Since the cost employed in misclassifying a bankruptcy company as non-bankrupt (Type I Error Region) or a non-bankruptcy company as bankrupt (Type II Error Region) will not be the same, the outcome of a model aiming to minimize the “costs” related to these error regions will be different from a model developed to minimize the “number” of companies in these regions. This study will try to answer the research question “Is the optimal cut off point considering to minimize the costs in Type I and Type II Error Regions in Bankruptcy Prediction equal to optimal cut off point considering to minimize the number of companies in Type I and Type II Error Regions?”

Another aspect that this study contributes is its focus on Turkish companies. As stated by Muzır and Çağlar (2009), although many models were developed about prediction of failure or bankruptcy, each model was designed by usage of samples from different countries. So, it is not clear that the known models will result the same levels of accuracies on Turkish companies or not, because there is a chance that the previous models and variables may fail in countries like Turkey with extreme levels of volatility in economic conditions. Hence, another aim of this study is to get a perspective about the previous models’ and variables’ application on Turkish companies.

1.2 Outline of the Study

Following the first section which summarizes the topic and the outline of the study, the second section presents the past literature about the usage areas of ratios. The review of past literature starts with an investigation about the usage of the ratios instead of the accounts building up the ratios. The predictive usage of ratios that had evolved by the help of different prediction methods and improvement of technology are deeply discussed. During the discussion, important topics like information theory or selection bias are also mentioned. Since the topic is related with the costs of bankruptcy, the methods to obtain these costs with empirical methods (rather than presenting a theoretical framework) are also studied.

In the third section, the selection and construction of the data and statistical method chosen are presented. In the fourth section, the model is constructed, verified and outputs of the model are discussed.

2. LITERATURE REVIEW

2.1 The Ground Work For the Usage of Financial Ratios in Academic Literature

By the beginning of 20th century, large amounts of financial data began to flow from many individual analysts, companies and institutions. To investigate, comprehend, compare and contrast them in better ways, financial ratios were developed by analysts.

Before starting to use the financial ratios in statistical methods, academicians called this usage into question,

- why these ratios were used instead of the financial accounts (creating the ratios) themselves
- if any assumptions were violated as the result of this usage or not.

Horrigan (1965) first investigated the statistical nature of financial ratios by analysing the data of a sample of companies from steel and petroleum industries between 1948 and 1957. (32 steel companies and 24 petroleum companies were sampled) . He concluded from his data that

- Financial ratios were not normally distributed or their dispersion was very large.
- Correlations between ratios existed (Multicollinearity).
- Correlation in single ratios over time existed (Autocorrelation).

Second, he evaluated the evidence on the utility of financial ratios, by making a detailed review of the past literature:

- Studies that were concerned with the ultimate form of financial difficulty: the failure of the firm,
- Studies that dealt with financial difficulties like defaults on debts and bank credit difficulties which were also troublesome but not as dramatic as business failures,
- Studies that dealt with a general financial condition like a lack of profits which could be indication of financial difficulties.

For these 3 groups of studies, he concluded that, Financial Ratios were very helpful instruments for reaching specific goals, and by the help of computers, a greater usage of unorganized data and more complicated statistical techniques would be possible to reach more accurate results.

Whittington (1980) mentioned that the empirical literature studied the statistical distributions, inter-relationships, and predictive contents of the accounting ratios. A common reason for using ratios was to reduce variables to similar scale to compare firms of different sizes. But, when the ratios were used, it was assumed that the ratio formation was the appropriate statistical form for summarising the data, without the consideration of what assumptions were necessary for this usage.

Proportionality was the assumption while using ratios. It was assumed that a proportionate relationship had to exist, between the two variables whose ratios were calculated.

One traditional use of ratios was to compare a ratio with some standard and observe if it was high or low. Another alternative use of them was the estimation of a functional relationship for the purposes of prediction.

But in an empirical relationship between accounting variables, the conditions necessary for proportionality might be violated in some situations:

- There might be a constant term in a relationship.
- The functional form of the relationship might be non-linear.

As an example, Gross Profit Ratio was given for the presence of a constant term which was commonly used in the analyses. If sales were the only source of gross profits, it would be reasonable to assume that zero sales would lead to zero profits. However, it was possible that because of the sources of income that were not related to sales, such as rents received, some Gross Profit might be obtained and this might result a constant term in the ratio.

Rather than comparing and contrasting the ratios, for estimating empirical relationships, regression analysis should be used in preference to ratio analysis. In regression analysis, the two conditions that violated the proportionality assumption were mitigated.

To investigate which of the ways was most valuable to represent the ratios in accounting studies (whether the comparison and contrast of the ratios, using them in linear regression or using in a non-linear regression model), return on net assets and the operating profit to sales were calculated by modelling with

- a simple proportionate relationship consistent with ratio analysis where a constant term was not present and the relation was assumed to be proportional,
- a linear relationship with a constant term,
- a quadratic term added, to test a very simple form of non-linear relationship.

The results of estimating the proportionate relationships showed that

- A high degree of association was present (indicated by R^2 in the regression, which was greater than 0.7 in each case, and a high value of the T- statistic associated with the estimated coefficient),
- Adding a constant term usually resulted a statistically significant estimate, suggesting that the ratio specification was inappropriate.
- Adding the quadratic term yielded a significant coefficient attaching to this term, suggesting that a quadratic specification was superior to either of the linear ones', but the introduction of the quadratic term led the constant term lose its statistical significance. (This constitutes a valuable framework for the further study that the ratios and nonlinear regression models can be used for analytical research.)
- The coefficient of the quadratic term was negative in the profits/net assets case (suggesting a downward curve) and positive in the operating profit/sales case (suggesting an upward curve).

Barnes (1987) analysed if the usage of ratios instead of absolute values in the accounting and finance literature was representative and created other problems or not. Like Horrigan, Barnes also commented that despite the extensive use of ratios, neither the conditions under which such use was appropriate nor the consequences of using these ratios when these conditions were not met were not discussed in the literature.

Instead of using the nominal values of accounting data, the reason ratios were used was to ensure the comparison by adjusting for size. In the same time, this brought the assumption that the financial ratios possessed the appropriate statistical properties for handling and summarising the data.

There were two reasons for the usage of ratios:

- To control the effect of size on the financial variables being examined. Size was properly controlled only when two financial variables were proportional. He commented that the violation of this proportionality didn't create too much trouble in literature, because, as financial ratios were constructed from two accounting variables, the joint distribution would depend on the behaviour of both the numerator and the denominator and on the relationship between these two. If there was non-proportionality, then the distribution of the ratio would be skewed, and this skewness was expected to lead some other separate insights and analysis.
- To control for industry-wide factors. Ratios helped the comparisons between a subject firm and its industry. In ratio analysis, a firm's ratios were compared with industry norms which might be a locational measure, such as the industry mean and median, and inferences about the firm's performance were based on the difference between the firm's ratios and the industry norm.

Barnes also commented on the normality requirement of the ratios. The normality of the ratios was required

- in order to compare the ratios between themselves (thinking of their effects on location measures),
- in order to facilitate the use of financial ratios in statistical models that assumed normality (The requirement of normality in order to facilitate some multivariate statistical model like MDA was the main motivation to develop other statistical methods like logistic regression that didn't require it in the data, as the literature about bankruptcy proceeded)

If present, 2 methods to overcome the non-normality were advised:

- The data might be applied a transformation that has a more normal characteristic. For example, square root and logarithmic transformations which resulted in normality of the data in most cases might be advised.
- Trimming the data by eliminating the outliers or changing an outlier's value to that of the closest non-outlier, and then attempting to fit the distribution to normal distribution.

Barnes' study concluded that, although creating some problems (like non-proportionality of the values used in the calculations of the ratios) ; when the supportive causes of using the ratios were thought (the convenience of using these ratios, the problems related with the usage of the absolute values, and the easy methods to overcome the problems arising by the usage of the ratios were considered), ratios would continue to constitute a powerful base for the analysis of the financial performance of a company or industry.

Kuh and Meyer (1955) stated that, when using the ratios in linear regression analysis, some conditions were to be fulfilled, to obtain the best linear unbiased estimates. They pointed out that, these conditions were frequently fulfilled with cross-sectional analysis, whereas time series analysis required more inspection. Indeed, most of the studies were using cross-sectional data as independent variables and there were little exceptional statistical models (i.e the hazard model) that use some panel data or time series data.

Besides the empirical implications of the financial ratios, Lev (1969) tested their practical usage by the companies and managers, with the hypothesis if firms had adjusted their financial ratios according to industry-wide averages or not. Managers would set the desired ratios as targets in their budgets and then regulate their operations such that the resulting

ratios would match with the budgeted ones. He tested his hypothesis with the Partial Adjustment Model.

$$\log y_{k,t} - \log y_{k,t-1} = \alpha + \beta (\log x_{k,t-1} - \log y_{k,t-1})$$

where

- $x_{k,t-1}$: The industry mean of the ratio which determines the target
- β : Coefficient or the speed of adjustment
- $y_{k,t}$: Value of a certain ratio at time t
- α : Constant Term (Expected to be 0)
- $t = 1$ to 20 (20 consecutive years were selected for each company)
- $k=1$ to 6 (6 different financial ratios were selected)

The data of 900 US companies between 1947 to 1966 were taken as the data set. The data were eliminated if a company had consecutive financial ratios in a period less than 20 years or if the belonged industry classifications included less than 10 companies. Here 6 financial ratios were chosen.

- Quick Ratio
- Current Ratio
- Equity to Total Debt
- Sales to Inventory
- Sales to Total Assets
- Net Operating Income to Total Assets

From the output of the linear regression,

- The value of β coefficient between 0 and 1 with high t-values was a strong indication that the firms periodically adjusted their ratios to the industry means.
- The α coefficient (constant) was found to be very small and statistically insignificant, leading to the conclusion that the model fitted the data well. If a significant α was present, this meant that, the companies adjusted their financial ratios not according to the industry averages.
- He also explained that β was higher for the ratios (like the current ratio) that involved short- term items and were under the direct control of management; therefore, they could be adjusted in the short run more easily (with less cost) than other ratios.

2.2 The Usage of Financial Ratios in Academic Literature

In a general perspective, the methods that financial ratios are used may vary in a wide range from

- descriptive methodologies to get status of a single company, companies, sector etc.
- certain decision making methods (like TOPSIS) to compare and contrast the companies and sectors within each other.
- statistical methods to make prediction about the financial position or performance of a single company.

Although each of these usage groups are summarized, as the topic of this dissertation study involves a predictive application of bankruptcy by using the ratios, the previous literature about predictive purposes will be discussed more deeply.

2.2.1 Descriptive Purposes

Ratios were popularly used to describe how similar or different the companies, industries or subsectors were.

Gupta and Huefner (1972) worked with industry-wide data to find the correspondence and relation between the industry attributes and accounting ratios. Rather than making inferences about the industries by using the ratios, they applied a cluster analysis to classify industries into a number of broad groups, by stating that the values of the ratios were expected to change from industry to industry.

They gathered data for the manufacturing companies, by categorising them within their SIC code groups. Their study had several important implications:

- The cluster analysis produced similar results with the classifications of the economists about industry-wide ratios.
- As the specialization of assets were greater, the link between the industries were also clearer. For example, “Fixed Assets” were assumed to be one of the most specialized assets in defining an industry, so the ratios related to fixed assets were distinctive between groups. On the other hand, as “Cash” was a highly generalized asset, the results about the distinctive ability of the ratios where cash was used was not clear.
- The inter-industry differences were disappearing when more aggregate ratios were used.

Gümüş, Köroğlu and Bulak (2018) analysed the financial ratios of 9 companies that were traded in Borsa Istanbul and represented in the BIST 30 Industrial Index (XUSIN). 21 ratios

to show liquidity, financial structure, activity and profitability positions were descriptively analysed from 2012 to 2016.

Ekşi and Akçi (2009) analysed the differences of financial ratios in subsectors of the manufacturing market. 26 ratios to show liquidity, financial structure, activity and profitability positions and market performance of the 158 companies were constructed in subsectors of manufacturing market (food & beverage, textile, chemical industry, mining, metal production, metal goods and machinery production, paper industry). Differences between the ratios of each sector were investigated with ANOVA test. As the result, although significant differences were observed in the ratios in mining and textile industry, the ratios of other subsectors were not different from the industry averages.

İçerli and Akkaya (2006) defined the financial failure as the announcement of negative profitability for 3 consecutive years, and selected 40 companies that failed and 40 that didn't fail in manufacturing sector between 1990 and 2003. They compared their 10 financial ratios with the Z Test. Current Ratio, Acid Ratio and Receivable Turnover were the ratios that were stated to be different with %95 confidence.

Coşkun and Sayılğan (2008) investigated the profitability ratios of the companies that had a financial failure and if the restructuring of the debts was a cure for the financial performance. Financial failure group was constructed with the companies, that had restructured their debt (19 companies) or had gone bankrupt without restructuring (1 company). In addition to these, the companies that had losses in 3 consecutive years or with negative equity were also added to the list.

The Gross Profit, EBIT and Net Profit accounts compared with the assets were analysed in years T_{-1} (1 year before the difficulty measures were observed), T_0 (the year that the difficulty measures were observed), T_1 (1 year after the difficulty measures were observed), Y_0 (the year that the debt was restructured) and Y_1 (1 year after the debt was restructured)

- The first hypothesis tested was if the profitability ratios in T_{-1} were different than the ratios in T_0 . The results showed that all of the profitability ratios deteriorated significantly by entering into the year of financial difficulty T_0 .
- The second hypothesis tested was if the profitability ratios in T_0 were different than the ratios in T_1 and T_2 , to indicate a recovery after the financial difficulty in the following years. The results showed that none of the profitability ratios were significantly different from the ratios of T_0 so this meant that no sign of recovery was observed in the following years.
- The third important hypothesis was tested between T_{-1} and Y_1 to observe if the companies reached their performance before difficulty, one year after the restructuring. It was observed that the profitability ratios were not different in T_{-1} and Y_1 , pointing out that after the restructuring these ratios had returned back to their values before the difficulty had arisen.

2.2.2 Decision Making Purposes & Methods

In some cases, the hypothesized question is not related to making a dichotomous decision about a single company if it will be bankrupt or not, but rather, sorting a number of firms to compare with each other. This comparison may be needed for an investment decision; an investor may be trying to find the best (or worst) performing company for a buying (or selling) decision. A manager may be trying to understand his company's financial position, with a comparison with other companies in the sector. For these purposes, statistical procedures may not be needed.

One of the ways to sort the companies by using several different attributes simultaneously is the TOPSIS (Technique for Order Preference by Similarity to the Ideal Solution) method developed by Hwang and Yoon in 1981.

TOPSIS method can be applied in 5 steps.

1. Calculation of normalized values by constructing a Decision Matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdot & a_{mn} \end{bmatrix}$$

Where m companies are to be compared by the help of n attributes.

$$r_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^m a_{ij}^2}} \quad \text{and} \quad R = \begin{bmatrix} r_{11} & r_{12} & \cdot & r_{1n} \\ r_{21} & r_{22} & \cdot & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdot & r_{mn} \end{bmatrix}$$

2. Calculation of weighted normalized values.

$$V_{ij} = W_j \times r_{ij}$$

3. Calculation of the distances with the best and worst conditions

$$d_{i\text{worst}} = \sqrt{\sum_{j=1}^n (V_{ij} - V_j^-)^2} \quad , \quad d_{i\text{best}} = \sqrt{\sum_{j=1}^n (V_{ij} - V_j^+)^2}$$

4. Calculating the distance to the Ideal Solution

$$C_m = \frac{d_{i\text{worst}}}{d_{i\text{worst}} + d_{i\text{best}}}$$

5. Sorting the companies according to the values of C_m

Sorting according to the TOPSIS method demonstrates the companies from the best choice to the worst. But the method lacks in 2 aspects.

- The method cannot underline any multivariate interaction that the statistical methods can.
- A statistical method is still needed to define the weights mentioned in step 2.

In their research, Türkmen and Çağıl (2012) analysed the financial performance of the 12 companies that were traded in Borsa Istanbul between 2007 and 2010 in Telecommunication sector. For each year, each company were evaluated according to the TOPSIS methodology and given a rank.

The 8 ratios used were Current Ratio, Acid Ratio, Receivable Turnover, Leverage, Sales to Total Assets, Net Profit Margin, Return on Equity and Return on Asset. The weigh matrix for these 8 ratios were defined respectively as

$$W = [0.14, 0.12, 0.13, 0.10, 0.15, 0.10, 0.12, 0.14]$$

From the ranking of the companies between 2007 to 2010, the financial performance trend of them were evaluated.

Likewise, Yurdakul and İç (2003) assessed the performance of 5 Turkish automotive firms and compared their results with companies' stock performance. 7 ratios they used were Net Profit Ratio (Net Income / Sales), EBITDA Margin, Gross Margin, Current Ratio, Inventory

Turnover, Efficiency (Sales /# number of employees), Current Liabilities / Total Liabilities, and the weights in their normalized matrix were respectively as

$$W = [0.135, 0.154, 0.135, 0.115, 0.135, 0.173, 0.154]$$

2.2.3 Predictive Methodologies Using Financial Ratios

2.2.3.1 Prediction of Failure or Bankruptcy

“Failure”, “Distress” or “Bankruptcy” (or “Non-going concern opinion” in the Audit Reports based studies) are the mostly used phenomena that are tried to be predicted using financial ratios. Although distress and bankruptcy are interrelated, they have some conceptual differences. Bankruptcy is one of the consequences of a distress. Rather than bankruptcy, distress can end with a turnover of the company or may lead to the reorganization of it. Bankruptcy is mostly related to the debt position of the company whereas distress may be related to the operations of the company and related with the income and liquidity. (Kulali, 2014)

Beaver’s descriptive study (1966) was a pioneer study to predict bankruptcy by the usage of financial ratios, in which financial performance of industrial companies were compared and contrasted.

In his research, a paired sample design was used to compare companies. 79 bankrupt companies (between 1954 – 1964) were paired with non-bankrupt companies within the same asset size and in the same industry.

When employing the paired sample design, he investigated the validity of it. While the paired design mitigated the disruptive influence of the industry and size factors, and provided a control over these factors, it might be eliminating any predictive power these factors might had led. Further in the study, he showed the correlation between the ratios and the sizes of the companies. He found that the ratios were uncorrelated with the asset sizes.

The financial statement data for the non-failed firms were obtained for the same fiscal years as those of their failed matching companies. 30 ratios were chosen according to their popularity in the literature, presence of a well performance in one of the previous studies, and the ratio to be defined in a cash flow concept. From these 30 ratios, 6 representative ratios were selected for the detailed analysis.

A main proposition was made: “The firm was a reservoir of liquid assets, which was supplied by inflows and drained by outflows. The reservoir served as a cushion or buffer against variations in the flows.”

Supporting sub propositions were also stated as

1. The larger the reservoir, the smaller the probability of failure
2. The larger the net liquid asset flow from operations, the lower the probability of failure
3. The larger the amount of debt, the greater the probability of failure
4. The larger the fund expenditures for operations, the greater the probability of failure

By the light of these sub propositions, how the values of these ratios should be changing in 2 classes of companies were assumed as in Table 2.1 below.

Table 2.1

Prediction of the Mean Values of Failed and Non-failed Firms

Ratio	Prediction
Cash Flow to Total Debt	Non-failed > Failed
Net Income to Total Asset	Non-failed > Failed
Total Debt to Total Assets	Failed > Non-failed
Net Working Capital to Total Assets	Non-failed > Failed
Current Ratio	Non-failed > Failed
No-credit Interval	Non-failed > Failed

Source: (Beaver W. , 1966)

In the Profile Analysis, the mean ratios of the companies were analysed for 5 years prior to failure. However, Beaver also stated that, without knowing the dispersion of the values, it was not possible to state if mean was a good predictor or not.

The Dichotomous Prediction Test was one of the first predictive attempts to predict bankruptcy before it occurred using financial ratios. To make predictions, the financial ratios were arrayed, each ratio was sorted in an ascending order and the array was inspected to find an optimal cut-off point. A cut-off point to maximize the percent of correct predictions was tried to be found. As the result of the Dichotomous Prediction Test, the values of the ratios to minimize the misclassification were found. Cash flow to total debt was found to be the ratio with the highest predictive ability, followed by net income to total debt and total debt to total

assets. So, he concluded that, from the propositions stated above, the second proposition, the net liquid asset flow from operations into the reservoir was the most, while the first proposition, the size of the reservoir was the least important factor in predicting and preventing the failure.

By the help of the Contingency Tables, Type I and Type II errors were analysed. Type I (the probability of misclassifying a failed firm) and Type II (the probability of misclassifying a non-failed firm as failing) errors were defined as the probabilities of error, conditional upon the actual status of the firm. His evidence indicated that the ratio analysis could not classify failed and non-failed firms with equal success. In each year before failure, Type I error was greater than the Type II errors. Type II error was remarkably stable over the five years while the Type I error increased as the time before failure increased. It was unfortunate that although the costs involved in the Type I error region were higher than the costs involved in the Type II error region, non-failed firms could be classified to a greater extent than the failed firms could be.

Finally Beaver stated that if ratios were used to detect the financial illness of a firm, there might be many firms whose illnesses were detected before failure occurred. In this case the proper treatment was applied and the firms didn't fail. The sample of failed firms would include those firms whose illnesses were not detectable through ratios. This created a biased sample for investigating the usefulness of ratios. An important piece of information missing was how many firms were saved from failure because their problems were detected in time through the use of ratios. However, this argument suggested that the study might only be understating the usefulness of the ratios.

Altman's study (1968) was a pioneer multivariate research in the bankruptcy prediction, using the financial ratios, in several aspects:

- The Z-Score Model developed by him constituted a benchmark to represent a certain level of accuracy achieved, and was mostly tried to be improved by the researchers of the topic in the following years.
- The financial ratios chosen were also relied for significant accuracy in prediction and commonly used by many researchers.
- Although possessing some assumptions, Multivariate Discriminant Analysis used in the study was the main multivariate statistical method trusted and led to many future research by the researchers.

Beaver's univariate approach was criticized: A firm with a poor profitability or solvency record might be regarded as a potential bankrupt. However, if it had an above the average liquidity, the situation might not be considered serious.

Altman analysed 33 bankrupt and 33 non-bankrupt companies in two groups. From initial 22 ratios as independent variables, he chose 5 ratios and applied multivariate discriminant analysis (MDA).

His discriminant function was

$$Z = 0.012 X_1 + 0.014 X_2 + 0.033 X_3 + 0.006 X_4 + 0.999 X_5$$

where

- X_1 : Net Working Capital / Total Assets
- X_2 : Retained Earnings / Total Assets
- X_3 : EBIT / Total Assets

- X_4 : Market Value of Equity / Book Value of Total Debt
- X_5 : Sales / Total Assets

Although having the highest coefficient, sales to total assets was the least significant ratio on an individual basis. The F score (that related the difference between the average values of the ratios in each group to the variability of values of the ratios within each group) was the lowest for this ratio. But it was still an important ratio. To find the contributions of the ratios to the discriminant function he calculated a scale factor which was computed by multiplying the variance of the ratio with the related coefficient. In the scaled factor, sales to total assets was the second ratio that contributed best.

1 year prior to bankruptcy, correctly classified number of companies were 94% of total. After the Z score was calculated from the discriminant function, it was seen that the companies with the Z scores greater than 2.99 were not bankrupt and the companies with Z scores lower than 1.81 were bankrupt in actual groups. And there was a grey area between 1.81 and 2.99 where all of the errors were made.

He stated that, the accuracy of the model decreased as the lead time increased. The prediction model was accurately forecasting up to 2 years prior to bankruptcy. All of the ratios showed a deteriorating trend as bankruptcy approached. And the most serious change in the ratios were seen between the third and the second years prior to bankruptcy.

Altman's study was very commonly criticized or taken as a benchmark in the following years. For instance, Joy and Tollefson (1975) critically analysed the discriminant analysis and its applications by Altman in bankruptcy studies. They also underlined the importance of the

prior probabilities of two dichotomous states (bankruptcy and non-bankruptcy) in building up the model. In the further research, the prior probabilities gained much more importance when expected costs related to bankruptcy prediction models were to be calculated.

“The Failing Company Doctrine” was the baseline of another study by Blum (1974). The Failing Company Doctrine stated “the possible harm to communities, employees, creditors, and owners caused by a failing business that might be forced into a liquidation, outweighed the harm to competition caused by allowing a failing and weak but still intact firm to merge with a competitor.” The reflection of this doctrine was also observed in the American Bankruptcy code, that favours the continuity of the entity to a preferable state rather than the ultimate bankruptcy of it.

Blum extended the Cash Flow Framework offered by Beaver that described the firm as a reservoir of financial resources and the probability of failure in terms of expected flows of those resources. Other things being equal, the probability of failure was more likely if

- the smaller the reservoir was (a larger reservoir would be a better buffer against uncertainties)
- the smaller the inflow of resources from operations in both the short and long-run was
- the larger the claims on the resources by creditors were
- the greater the outflow of resources required by the operation of the business was
- the more variable earnings and claims against resources were (represented both by outflows to maintain current operations and by obligations to creditors) ; the less variable were inflows and outflows, the more likely that future events could be predicted.
- the more failure-prone the industry locations of a firm's business activities were (certain industries at specified times- such as automobile manufacturing in the early twentieth

century or prefabricated home construction in the early 1960s- were more failure-prone than other industries.)

Representing the items in Cash Flow Framework, the failing company model was constructed with reference to three groups: liquidity, profitability, and variability.

THE FAILING COMPANY MODEL

I. LIQUIDITY

a. Short-Run Liquidity

Flow	:	1. The quick flow ratio
Position	:	2. Net quick assets/inventory

b. Long-Run Liquidity

Flow	:	3. Cash flow/total liabilities
Position	:	4. Net worth at fair market value/total liabilities
		5. Net worth at book value/total liabilities

II. PROFITABILITY : 6. Rate of return to common stockholders who invest for a minimum of three years

III. VARIABILITY : 7. Standard deviation of net income over a period
8. Trend breaks for net income
9. Slope for net income
10-12. Standard deviation, and slope of the ratio,
net quick assets to inventory

Quick Flow ratio was defined as $\text{Cash} + \text{Notes Receivable} + \text{Market Securities} + (\text{Annual Sales} \div 12) \div (\text{Cost of Goods Sold} - \text{Depreciation Expense} + \text{Selling and Administrative Expense} + \text{Interest}) \div 12$.

- The quick flow ratio related reservoir size and resource inflow to resource outflow.
- Net quick assets/inventory indicated the relationship of both current liabilities (short-term resource claims) and inventory (relatively illiquid component of the reservoir) to the highly liquid quick assets (cash and equivalents, plus accounts and notes receivable).
- Cash flow/total liabilities related resource inflow to total claims;
- Net worth measures related reservoir size to total claims.
- The profitability measure, rate of return to common stockholders, reflected all of the elements of the cash-flow framework.
- The last six variables indicated the variability and trend of resource inflow (net income) and of the short-term liquidity indicator, net quick assets/inventory.

Multiple Discriminant Analysis was used in The Failing Company Model, and it predicted failed and non-failed companies with an accuracy of %93 to 95 at the first year before failure. Predictive accuracy was approximately %80 at the second year before failure and %70 at the third, fourth, and fifth years before failure. Discrimination between failed and non-failed companies was not found to be statistically significant at the sixth year before failure.

Besides making prediction about the failure of companies, Blum's study had further implications:

A Model without Financial Ratios: Since discriminant analysis assigned weights, it might assign weights to absolute accounting data rather than ratios. An experiment was conducted to test the predictive accuracy of a non-ratio model. Five of the ratios used in the Failing Company Model were broken down into their component numerators and denominators. The total of eight different numerators and denominators were used as eight independent variables

in conjunction with the four non-ratio variables of the model to form a non-ratio version of the Failing Company Model. Although The non-ratio model was not as accurate as the Failing Company Model in the first year before failure, it was more accurate thereafter (for ranges with five-year intervals).

Anticipation of Failure by the Market: In order to indicate stock market's ability to anticipate the timing of failure, a measure of "Going to the Market" was calculated by the fair market value of the shares offered to the public divided by the sum of the fair market value of net worth of the company at the end of each of the years in the range.

If the market anticipated failure accurately, it was expected that the market would not invest further in a failing firm and Going to the Market would be higher for non-failing firms than failing firms. But the results showed the market preferred to invest in failing rather than non-failing firms.

Comparison of the Failing Company Model with Altman's multivariate study, Blum claimed that although Altman's model was predicting failure with an accuracy of 94% prior to bankruptcy year on the primary sample and 96% on a validation sample; after second year, the model produced illogical results. Beyond the third year, Altman's model became less accurate than even a chance procedure. In the fourth year, for example, accuracy could be increased to 71% by inverting the decision rule as: When failure would have been predicted, predict nonfailure, and vice versa. However, at the time of decision it was impossible to know whether the firm was at the fourth or second year before failure.

Poyraz and Uçma (2006) evaluated the Z-Scores of the tourism, agriculture, textile and transportation sector ratios from 1992 to 2003. The aim of the usage of Z-Score with the sector ratios was to analyse if the crisis years of 1994 and 2001 were producing low Z-Scores or not. They concluded that, Altman Z-Score was not a good indicator of financial crisis for the mentioned industries.

Terzi (2011) evaluated the Turkish Food companies according to their Z-Scores in 2009 and 2010. He categorized the companies as successful and unsuccessful, and tried to derive the distinctive financial ratios between these 2 groups. Without mentioning a predictive ability of the ratios, debt ratio, equity ratio, net working capital to total assets, EBIT to Interest Expenses, Profitability Ratio and Return on Assets were seemed to have discriminating ability between successful and unsuccessful companies.

2.2.3.2 Prediction of Other Phenomena

Financial Ratios were frequently used in prediction of other phenomena rather than the Bankruptcy.

Oktar and Dalyancı (2010) developed a probit model to predict the financial crises of Turkish Economy, using the national aggregate data and economic indicators. Their model concluded that the decrease in Turkish Central Bank (TCB) Composite Leading Indicator for Turkish Economy, deficit in the current account balance, decrease in national reserves and decrease in industrial production, increases in TRY/USD parity and BIST100 index contributed significantly to the prediction of a financial crisis in Turkish Economy.

Casey et al. (1986) tried to make a discrimination between the unsuccessful companies in the groups of reorganized and liquidated. Using a sample of 113 companies (57 reorganized and 56 liquidated) and by using a probit regression, they showed that there were significant independent variables that had discriminated the reorganized companies from liquidated. The discriminating independent variables were the change in the total assets, retained earnings/total assets and free assets percentage (free assets were the assets that were not secured by previous borrowing) . The model correctly classified %70.8 of the companies in correct groups.

Rather than prediction of ultimate failure, Lau analysed the transitional states, that started from a healthy firm to a state of bankruptcy (1987).

Lau defined five financial states as

- State 0: Financial Stability: 350 companies were selected which were known to experience no financial distress or loss between 1972 – 1977
- State 1: Ommiting or Reducing Payments: The companies that had a stable dividend payment record during 1972 – 1975 and declining trend in payments from 1976 – 1978. 20 companies were selected for training sample.
- State 2: Technical default and default on loan payments. The companies that were reported from SEC 10-K reports to default loan interest and/or principal payments during 1976 and 1977. 15 such companies were selected for training sample.
- State 3: Protected under Chapter 10 or 11 of the Bankruptcy Act. 10 Companies were selected.
- State 4: Bankruptcy and Liquidation: 5 such companies were selected.

10 independent variables were constructed, showing the borrowing capacity, stock flexibility, cost flexibility, dividend flexibility and asset disposability of the companies. The mean and median values of these variables showed that there were significant differences in values among different state companies, and there were no multicollinearities between the variables.

From the variables, a “Multinomial Logit Analysis” (MLA) was constructed as

$$Z_j = b_{j1} X_1 + b_{j2} X_2 + \dots + b_{j10} X_{10}$$

$$p_j = \exp(Z_j) / \sum_{j=1}^J \exp(Z_j)$$

where

- p_j : The probability of a given firm to enter state j

The MLA had several distinctive evaluation criteria:

For example, when the actual state for a company was state 1, let's think that the MLA produced an output as

- Output A produced a probability of %50 for entering into state 0 and %0, %50, %0 and %0 probabilities for the other 4 states respectively $F_A = (\%50, \%0, \%50, \%0, \%0)$
- Output B produced a probability of %50 for entering into state 0 and %0, %0, %0 and %50 probabilities for tother 4 states respectively $F_B = (\%50, \%0, \%0, \%0, \%50)$

A scoring rule to that ignored the ranking of the states would give F_A and F_B the same score.

However, if state 1 was considered closer to state 3 than state 5, then F_A was better than F_B .

Ranked Probability Scoring Rule was applied to solve this problem and it produced an evaluation score of 0.875 for F_A and 0.75 for F_B .

So as the result of the MLA and Ranked Probability Scoring, the classification accuracies in different states were as given in the Table 2.2 below.

Table 2.2

Accuracies of the Different Financial State Estimations of MLA

State 0	%99,4
State 1	%65,0
State 2	%86,7
State 3	%70,0
State 4	%60,7

Source: (Lau, 1987)

Although the accuracies especially in the bankruptcy states (state 3 and 4) were low when other studies were considered, Lau's analysis were important to classify the companies into the transitional stages before bankruptcy. The outcomes of the study can be applied for the companies to get the early warnings of a bankruptcy and have a Turnaround.

Poston et al. tried to predict Turnaround (1994). They studied 204 distressed companies during the period from 1970 to 1976. Each company was evaluated eight years subsequent to the year of entry to their sample and assigned to one of three groups according to its financial condition at that point in time.

Their main contribution with their study was the criticism about the utilization of a choice based sampling technique in the past literature to identify the financial ratios that discriminate among healthy and unhealthy firms. Like Zmijewski (1984), they pointed out that this

methodology of oversampling distressed company resulted in estimation bias. This sampling methodology ignored the possibility that a firm which exhibited potential for failure at one point in time might reverse its negative trend before failure actually occurred. In other words, prior research ignored the corporate turnaround phenomenon, because of the choice-based sampling.

Fitzpatrick (1934) defined the transitional stage of a business failure as

1. The incubation
2. Financial embarrassment
3. Financial insolvency
4. Total insolvency
5. Confirmed insolvency.

Avoiding a choice-based sample and representing 5 stages of distress of Fitzpatrick, Poston et al. included the companies into their data if they met any one of the following three criteria:

OPLOSS: Two or more consecutive operating losses. It was an indicator of weakened short-term earnings power. Firms entering this sample could be in the incubation stage.

CRLESS: A current ratio less than 1.0 as of the end of any single fiscal year. Firms in this group were in financial embarrassment or financial insolvency stage.

NEGRET: A negative retained earnings account as of the end of any fiscal year. Companies in this stage were assumed to be in the later stages of business failure because they had indicators of impaired long-term earnings power.

In the next step, each company was assigned a value for the categorical dependent variable, STATUS. If a company filed for bankruptcy or liquidated at any time during the eight-year period, it was classified as a business failure, and assigned a value of 1 for STATUS. If, after 8 years, the company was still in financial distress, according to the three criteria listed above, it was classified as a survivor with unknown ability for continued success and assigned a value of 2 for STATUS. If a company was no longer financially distressed, based on the same three criteria, it was classified as a turnaround and assigned a value of 3 for STATUS. Each 204 firms' financial condition was traced and evaluated 8 years subsequent to the year of entry into the sample. 46 companies were classified as failures (STATUS 1), 123 firms were classified as Turnarounds (STATUS 3) and 35 companies were in existence but still met at least one of the three entry criteria (survivors of a failure with unknown ability).

After grouping the companies into three groups as dependent variable, they applied 2 different models for prediction of turnaround: Z-score model and Probit Model (with seven ratios: Income/Total Capital, Sales / Net Plant, Inventory/Sales, Total Liabilities/Total Capital, Receivables/Inventory, Quick Assets/Current Liabilities, Cash/Total Assets). None of these models gave good prediction about distressed company that achieved turnarounds. But there were differences in the results the models introduced. The Z score model was biased toward defining distressed companies as failures. They demonstrated a high Type II error rate. Whereas the Probit Model was biased toward classifying failing companies as Turnarounds, resulting high Type I error rates

Kalaycı and Karataş (2005) tried to develop a model to predict the stock returns from the financial ratios. They developed their model in 2 ways. First, they constructed factors from the ratios by Principle Component Analysis and made Regression Analysis using these factors. Second, they made the Regression Analysis by using the ratios themselves. Food and beverages, manufacture of chemical, petroleum and plastics, and manufacture of paper and paper products industries were analysed for 1996 and 1997. Although low adjusted R^2 values were found from each regression analysis, the independent variables which were significant in explaining the returns were noted. And it was concluded that stock returns were explained by profitability, stock market performance, and productivity ratios for the above sectors in related periods.

One of the areas that the bankruptcy prediction studies serve is the area of Auditing. Not necessarily predicting an ultimate bankruptcy, these studies help the auditors in decision making between the types of the qualifications they conclude, because the auditors' qualifications are systematically related to financial variables.

Auditors need aid in decision making because their decisions involve economic trade-offs between qualifying too frequently (and risk losing the clients) and qualifying rarely (and being exposed to lawsuits and losing reputation). Dopuch et al. (1987) developed a model with a sample of 275 first-time qualifications and 441 clean qualifications between 1969 – 1980. The qualifications were divided into

- Going Concern: The auditor expressed doubt about firm's ability to finance future activities
- Litigation: If a lawsuit against the firm created an uncertainty about the future
- Asset Realizing: If the realizable value of an asset was less than its book value

- Multiple of above

These qualified opinions indicated a warning signal about the receipt of another qualified opinion later on, because according to the study, %75 of the companies received another qualified opinion in the subsequent year.

Taking into account that these were the first time qualifications and the previous qualification was a clean opinion (for the qualified group), Dopuch et al. investigated if the changes in the financial statement or market variables would capture a change in a state that was associated with the change in the audit opinion from clean to qualified.

The variables used in the WESML (Weighted Exogeneous Sample Maximum Likelihood) Probit Model were

$\Delta \frac{Total Liabilities}{Total Assets}$	Δ Residual Standard Deviation of Returns
$\Delta \frac{Receivables}{Total Assets}$	Returns – Industry returns
$\Delta \frac{Inventory}{Total Assets}$	Book Value of Assets
Δ BETA	Current Year Loss
	Time Listed

The output of the probit model provided evidence that, by using the publicly available financial and market data, it was possible to predict if an auditor would issue a first time qualification in the current year or not. The probabilities of the probit model also showed evidence about the type of the qualified opinion, Going Concern opinions have the highest score, Litigation opinions have the lowest and Asset Realising having an intermediate score.

As Dopuch et al. tried to predict the Audit Opinion from financial data, Hopwood et al.

(1989) tried to predict the bankruptcy from Audit Opinion. They searched if any kind of a qualified opinion had a predictive power in predicting bankruptcy and could be counted as a red flag for bankruptcy before it happened.

64 companies that went bankrupt between 1974 to 1981 and a control group of 55 non bankrupt of the same time were selected.

With the Log-Linear Analysis, they constructed the model with financial variables and auditor opinion variables as

$$\ln \frac{p(\text{Bankruptcy}|\text{Ratios\&Opinion})}{p(\text{Nonbankruptcy}|\text{Ratios\&Opinion})} = \mathbf{a} + \mathbf{h_1R_1} + \mathbf{h_2R_2} + \mathbf{h_3R_3} + \mathbf{h_4R_4} + \mathbf{h_5R_5} + \mathbf{h_6R_6} + \mathbf{h_7C} + \mathbf{h_8ST} + \mathbf{h_9G}$$

where the coefficients and related independent variables were

- h_1 :Net Income / Total Assets
- h_2 :Current Assets / Sales
- h_3 :Current Ratio
- h_4 :Current Assets / Total Assets
- h_5 :Cash / Total Assets
- h_6 :Long Term Debt / Total Assets
- h_7 :Consistency opinion dummy variable
- h_8 :Subject-to opinion dummy variable
- h_9 :Going Concern opinion dummy variable

To express the predictive ability of the loglinear models, R was calculated as

$$R^2 = (\text{model } \chi^2 - 2p) / -2L(0)$$

where

- p : the number of variables in the model excluding the intercept
- $L(0)$: the maximum log likelihood with only an intercept in the model

They demonstrated the output of two loglinear models -with and without the auditor opinions- and calculated the additive information and predictive power of these opinions.

The analysis had some important implications. Table 2.3 and Table 2.4 below show that both models with or without the auditors' opinion had strong predictive capability in predicting the bankruptcy one year prior to bankruptcy (with $R=0.65$ and 0.69 in models respectively) and the incremental χ^2 values on Table 2.4 show that the entrance of the auditor opinions into the model significantly increased the predictive power.

Table 2.3

Ratios Only Model

$$\ln \frac{p(\text{Bankruptcy}|\text{Ratios})}{p(\text{Nonbankruptcy}|\text{Ratios})} = a + h_1 R_1 + h_2 R_2 + h_3 R_3 + h_4 R_4 + h_5 R_5 + h_6 R_6$$

	Years Before Bankruptcy				
	5	4	3	2	1
Model R	<u>0.46</u>	<u>0.41</u>	<u>0.41</u>	<u>0.57</u>	<u>0.65</u>
Coefficients:					
a	0.25	0.49	-1.72	-2.88	-0.80
h ₁	-36.62	<u>-30.43</u>	<u>-16.59</u>	<u>-20.85</u>	<u>-32.36</u>
h ₂	2.22	1.57	2.80	2.81	0.55
h ₃	<u>-0.49</u>	<u>-0.32</u>	<u>-0.44</u>	<u>-0.57</u>	<u>-0.07</u>
h ₄	1.89	1.13	2.03	3.44	1.42
h ₅	7.27	5.68	3.23	5.94	-7.45
h ₆	-0.22	-0.62	3.92	5.47	2.70

Significant at %1 level**Significant at %5 level**

Significant at %10 level

Source: (Dopuch, Holthausen, & Leftwich, 1987)

Table 2.4

Ratios and Auditor Opinion Model

$$\ln \frac{p(\text{Bankruptcy}|\text{Ratios\&Opinion})}{p(\text{Nonbankruptcy}|\text{Ratios\&Opinion})} = a + h_1R_1 + h_2R_2 + h_3R_3 + h_4R_4 + h_5R_5 + h_6R_6 + h_7C + h_8ST + h_9G$$

		Years Before Bankruptcy				
		5	4	3	2	1
Model R		<u>0.41</u>	<u>0.41</u>	<u>0.43</u>	<u>0.56</u>	<u>0.69</u>
Incremental χ^2 to above model		0.41	5.38	8.25	5.45	<u>15.87</u>
Coefficients:						
a		0.20	0.55	-2.06	<u>-3.91</u>	-3.28
h ₁		<u>-35.61</u>	<u>-24.90</u>	-12.51	<u>-20.58</u>	<u>-30.84</u>
h ₂		2.23	1.19	2.92	2.50	-1.66
h ₃		-0.50	<u>-0.88</u>	<u>-0.95</u>	<u>-0.54</u>	0.04
h ₄		1.93	2.63	3.53	<u>4.80</u>	5.13
h ₅		6.76	7.67	4.18	5.52	-10.99
h ₆		-0.45	-0.02	4.82	6.05	4.62
h ₇		0.16	-0.05	0.33	1.08	2.50
h ₈		0.54	-0.30	0.70	-0.10	1.02
h ₉		1.38	22.97	21.18	7.51	<u>11.40</u>

Significant at %1 level**Significant at %5 level**

Significant at %10 level

Source: (Dopuch, Holthausen, & Leftwich, 1987)

Koh and Killough (1990) developed a model to aid the going concern opinions of the auditors, and tested their model against auditors' opinions in bankruptcy prediction. They chose 35 bankrupt companies that filed bankruptcy between 1980 and 1985 to construct their stepwise discriminants analysis. Using 21 independent variables, the discriminant function was built up as

$$Z = -1.2601 + 0.8701 X_1 + 2.1981 X_2 + 0.1184 X_3 + 0.8960 X_4$$

where

- X_1 : Quick ratio
- X_2 : Retained Earnings to Total Assets
- X_3 : Earnings per Share
- X_4 : Dividends per Share

Choosing the critical cut-off value as 0, the model had a classification accuracy of %92.65.

In the second step, they tested their model against the unqualified opinions of auditors prior to bankruptcy. 400 companies were chosen randomly, of which 14 were bankrupt. They assumed that, there should be a qualified opinion prior to the bankruptcy. Of 14 companies, only three were given any kind of qualified opinion (%21.43) whereas their model was pointing out 11 companies again as bankrupt. (%78.57). So, the model was highly accurate when the auditor opinions prior to bankruptcy was compared.

2.2.4 Predictive Methodologies Using Market Variables

Although admitting the success of the Ratio Based Models in predicting bankruptcy, Aharony et al. (1980) underlined that these models had no theoretical foundations. They said that the market driven data had received little attention in studies of corporate bankruptcy. So, they tried to predict corporate bankruptcy by using the market data as the predictor.

Their sample was 45 industrial bankrupt companies that had gone bankrupt between 1970 – 1978 and their controlled non-bankrupt partners. Weekly rates for each company's returns and market return were calculated.

From the CAPM equation

$$\text{Var} (R_j) = \beta_j^2 \text{Var} (R_m) + \text{Var} (e_j)$$

where

- $\text{Var} (R_j)$: total variance of returns of security j
- $\text{Var} (R_m)$: variance of returns of the market portfolio
- $\text{Var} (e_j)$: variance of the disturbance term of security j

In Cross-sectional analysis, the null hypothesis that “the means of the total variance $\text{Var}(R)$ for the bankrupt and control groups were equal at any given point in time relative to the event week” was tested.

Table 2.5

Total Variance of Returns for Failed and Non-failed Firms

Week before bankruptcy	Bankrupt Group Var(R)	Nonbankrupt Group Var(R)	T value of Differences
1	0.0117	0.0046	<u>6.75</u>
20	0.0101	0.0045	<u>5.68</u>
52	0.0091	0.0046	<u>4.35</u>
104	0.0075	0.0044	<u>4.04</u>
156	0.0065	0.0040	<u>4.07</u>
208	0.0055	0.0039	2.84
260	0.0054	0.0038	2.58
312	0.0052	0.0035	2.20
350	0.0050	0.0032	2.66
390	0.0050	0.0029	2.55

Significant at %0.05 level (One tailed test)**Significant at %0.5 level (One tailed test)****Significant at %1 level (One tailed test)***Source: (Aharony, Jones, & Swary, 1980)*

As shown in Table 2.5 above, the total variance of returns for the bankrupt group was significantly greater than the non-bankrupt group and the level of significance was increasing as the bankruptcy approached. More deeply analysed, it was observed that the differences were not due to market wide, but company specific factors.

Clark and Weinstein (1983) studied the effects of a bankruptcy announcement and the signals of a bankruptcy to the market performance of the related stocks. They measured these signals with several performance indicators like:

$$AR_t = \sum_{i=1}^{N_t} R_{it}$$

$$CAR_K = \sum_{t=T}^K AR_t$$

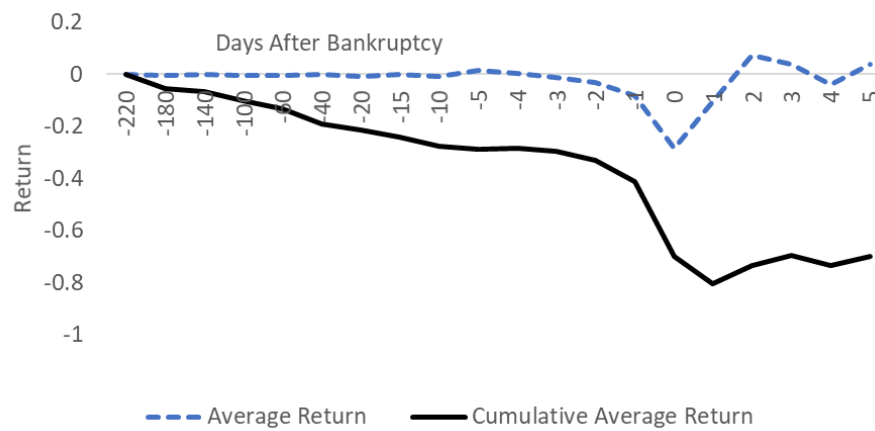
where

- AR_t : The Average Return on the related day
- R_{it} : Market adjusted return on stock I during related day
- N_t : Number of bankrupt stocks in related day
- CAR_K : Cumulative Return until day K (Starting from day -220)

When the Average and Cumulative returns of bankrupt stocks were examined as shown in the Figure 2.1 below, it was seen that the shareholders burdened larger losses during the month the bankruptcy had occurred and those losses were concentrated in three trading day interval around bankruptcy filing date.

The quick and sharp decrease and the continuing losses after the bankruptcy filing date showed that, the filing event conveyed information to the market that was not anticipated before the event.

Figure 2.1

Daily and Cumulative Returns Before and After Bankruptcy Filing

Beaver (1968) also made a comparison about the effect of a bankruptcy on the market returns by analysing 79 failed and 79 non-failed firms of the period from 1954 to 1964.

$$R_{it} = \frac{D_{it} + P_{it} - P_{it-1}}{P_{it-1}}$$

where

- P_{it} : Price for security i at time t
- P_{it-1} : Price for security i at time $t-1$ (adjusted for changes like stock dividends or splits)
- D_{it} : Cash dividend paid on security i between time $t-1$ and t

The return analysis by Beaver made clear that the median value of the failed companies' returns was below the non-failed, and the difference in medians increased as failure had approached.

To find whether the market or the ratios predicted failure before, Beaver made an assessment as

- The time in which the returns had a sharp decline so large that very few non-failed firms ever had was defined as the time the market predicted failure. This sharp decline was formulated as $R_{it} = -0.50$ and $e_{it} = -0.60$
(e_{it} is the residual rate of return for security i from time $t-1$ to t ; $e_{it} = R_{it} - R_{mt}$)
- The time in which the cash flow to total debt ratio fell below 0.02 and never rose above 0.08 subsequently was defined as the time the cash flow to total debt ratio predicted failure.
- The time in which the net income to total assets ratio was -0.01 and never rose above 0.05 subsequently was defined as the time the net income to total assets ratio predicted failure.
- The time in which the total debt to total assets ratio was 0.65 and never fell below 0.45 subsequently was defined as the time the total debt to total assets ratio predicted failure.
- The time in which net working capital to total assets ratio was 0.10 and never rose above 0.30 subsequently was defined as the time the net income to total assets ratio predicted failure.

So, for 29 bankrupt companies for which both rate of return data and the ratios data were obtainable, the related time were noted and compared. Using these three events above as signals, it was seen that the market returns data anticipated failure first. As for the timings of bankruptcy signals, the market returns were giving signals 2.45 years before failure on average and as for the other ratios, the signals were produced 2.17, 2.31 . 1.45 and 0.97 years for cash flow to total debt, net income to total assets, total debt to total assets, net working capital to total assets ratios respectively.

In addition, by using the above signals as predictors of failure, the greatest accuracy was reached by the net income to total assets ratio, capturing 49 of the failed firms correctly out of 53 firms.

2.3 Development of Literature about Prediction of Failure

The predictive research employing the financial ratios in univariate analysis (pioneered by Beaver) and then multivariate analysis (pioneered by Altman) was needed to be developed by certain groups of other modern researchers because

- In using the financial ratios as independent variables and making predictions, the research mostly lacked the theoretical background in constructing the model and choosing the independent variables; although there were some attempts like Wilcox (1971) to explain the effects of ratios on failure by leaning on a theoretical base. The empirical findings of the research were sample specific, and not permitting generalized results, over different samples.

Supported by important number of researchers, some important research about bankruptcy prediction was directed using Cash Based Funds Flow Models.

On another advancing pathway, some group of researchers made progress in other statistical ways to group the companies dichotomously using Probabilistic Regression Techniques; originating from the proposition that “the multivariate analysis implied too much statistical assumption, that were hard to realize in the real world conditions”. These techniques will be analysed in detail.

2.3.1 Prediction of Failure with Cash Flow Basis Accounts

Cash Based Funds Flow Framework researchers tried to improve the classification results in predicting the bankruptcy by building up the empirical study onto a robust theoretical framework of Cash Flow Basis at the same time.

Casey and Bartczak (1985) tried to improve Altman's 1968 model that used 6 accrual basis accounting independent variables.

- Cash/Total Asset
- Current Assets/Total Assets
- Current Assets/Current Liabilities
- Sales/Current Assets
- Net Income/Total Assets
- Total Liabilities/Owners' Equity

Adding 3 more Cash Flow Basis Accounting Independent Ratios, they tried to improve the bankruptcy model:

- Cash Flow from Operations (CFO): Net Working Capital provided by operations, (+ / -) changes in the non-cash working capital accounts except for short-term indebtedness.
- CFO / Current Liabilities (CFCL)
- CFO / Total Liabilities (CFTL)

Using these 6 accrual basis accounting independent variables plus one or more cash flow basis accounting variables in both MDA and Stepwise Logistic Regression, improvement of the previous models was aimed.

60 of 105 companies filed for bankruptcy between 1971 – 1982 were selected for the bankrupt group. The cash flow data were collected by requiring all audit opinions for the first year data to be dated prior to the bankruptcy filing date, complying with the data timing requirements mentioned in Ohlson 1980 study. 230 non-failed firms were chosen from COMPUSTAT to match the industrial classifications of the failed firms. The companies were not matched according to their sizes.

In the conclusion of their study, although the models constructed by the operating cash flow variables were producing significant results, they were not able to improve the classification accuracy that was achieved by accrual basis ratios, neither with MDA nor with probit. But it was commented that, some other possible uses of operating cash flow variables might be explored in further research. For example, although these variables didn't improve accuracy in bankruptcy prediction, they might be helpful in explaining other phenomena like corporate acquisition, loan defaults and dividend omissions.

Gentry et al. (1985) made a pioneer study in the area of bankruptcy prediction with the funds flow components. In fact, their research was a continuum of the results of Casey and Bartczak research. They tried to determine if the funds flow components by themselves were able to discriminate between failed and non-failed companies. They redesigned the funds flow model with

- Funds from Operations (NOFF)
- Working capital Funds Flows (NWCFF),
- Financial Funds Flow (NFFF)
- Fixed Coverage Expenses (FCE)
- Capital Expenditures (NIFF)

- Dividends (DIV)
- Other Asset and Liability Flows (NOA&LF)
- Change in Cash and Marketable Securities (CC)

Their model was

$$\text{NOFF} + \text{NWCCF} + \text{NFFF} + \text{FCEF} + \text{NIFF} + \text{DIV} + \text{NOA\&LF} - \text{CC} = 0$$

Since total net inflows should equal total net outflows during a certain amount of time.

Each of these variables were divided by the total net flows (TNF) and to model the effects of these in and outflows, the variable CC/TNF was omitted.

114 companies were selected as the deleted companies from the COMPUSTAT records between 1970 – 1981. But bankruptcy was not the only reason for deletion from these records. From these 114 companies, the companies that were liquidated, declared bankruptcy were selected and 92 companies were selected as failed companies (Other reasons of deletion were selling of assets to another company, mergers& acquisition, stopped trading, etc.). From these 92 companies, only 33 of them had financial ratios for 1, 2, and 3 years before failure. Each of the failed companies were matched according to industry, size and complete data criteria.

Like Casey and Bartczak, their cash flow based variables were not able to improve the classification results. But one major implication of their study was that dividend funds flow component was a significant variable in the classification. The smaller the relative dividend component, the higher the probability of failure was. That was probably because, a failing

firm would experience a shortfall of funds from the operations, and causing a reduction of dividend payments.

The contribution of the cash flow based ratios to the prediction of failure was also tried to be proven by Gombola et al. (1987) . 24 ratios, of which 9 were cash flow based were chosen first. These were

- Income Plus Depreciation (INCDEP)
- Working Capital from Operations (WCFO)
- Cash Flow from Operations (CFFO)

Each were divided by sales, assets and debt.

77 failed companies were matched with the non-failed partners according to the size and industry between 1967 – 1981.

The analysis was divided into two separate time divisions. 1967 – 1972 and 1973 – 1981, that was mostly overlapping with the interval used by Casey and Bartczak (1985) and Gentry et al. (1985) data. This was done to show that if Cash Based Funds were significant predictors of the bankruptcy after the mid-1970s and on, and probably not before.

Factor analysis were used to analyse how these 24 variables were grouped.

- Six factors were observed for the first period: cash position, current assets to sales or assets, current debts to total debts, debt position, return on assets, and return on sales.
- For the second period, seven factors were observed as: cash position, current assets to sales or assets, current debts to total debts, debt position, return on assets, cash flow, and return on sales.

The main difference in the groupings was that cash flow factor was appearing as a separate factor in the second period, although it was not present as a separate factor in the first period. Out of 24 ratios, the ratios with the highest load were chosen for entry into the MDA model. Again, four separate models were developed. In the first MDA trial, none of the Cash Flow Based variables were included. In the second, CFFO ratios were included, in the third and fourth, INCDEP and WCTO ratios were added respectively.

As the result of their analysis, in parallel with Casey and Bartczak (1985), the cash flow based ratios were not able to marginally improve the classification accuracy achieved by accrual basis ratios.

Dambolena and Shulman (1988) tried to obtain a more accurate model by defining a “Net Liquid Balance” as a new independent variable. Net Liquid Balance (NLB) was the difference between Liquid Financial Assets (i.e., cash, marketable securities) and callable liabilities (i.e. short term debt, current maturity due on the long term debt). 50 failed, and 50 non-failed companies were analysed with stepwise logistic regression, and the inclusion of NLB to the Gentry et al. (1985) independent variables improved the predictive performance prior to one and two years of bankruptcy.

Cash flow models were mainly differing in how they formulate the cash flow of the company. Van Frederikslust (2012) defined the failure in the “Cash Inefficiency” model. Cash-based funds flow model developed by Helfert(1982) were used in the research by Gentry et al. (1985). Helfert(1982) model was mostly helpful to provide analysis by the area of management attention. In “Bankruptcy Prediction-An Investigation of Cash Flow Based

Models” (Aziz, Emanuel, & Lawson, 1988), Cash Flow Identity model developed by Lawson (1971) was used, which was mostly related with the firm valuation.

Van Frederikslust defined the failure as the moment when the cash flow from operations plus the proceeds from new loans and liquidation of assets plus the opening balance of liquid reserves was not enough to pay the obligations. The reason for failure was mostly a decline in sales, leading to a decrease in cash inflow from operations. A decline in sales was mostly preceded by a recession, the loss of an important customer, shortage of raw material, deficiencies of management, etc.

So, the moment of failure was at time t when the Cash Balance $CB_t < 0$

$$CB_t = CB_{t-1} + S_t - (AR_t - AR_{t-1}) + OI_t - GS_t - GA_t - (BI_t - BI_{t-1}) + (SD_t - SD_{t-1}) + ALD_t$$

where

- CB_t : Cash Balance at the end of period t
- S_t : Sales in period t
- AR_t : Accounts Receivable (including prepaid expenses) in period t
- OI_t : Other Income (including liquidation of assets) in period t
- GS_t : cost of Goods Sold in period t
- GA_t : General, Administrative (including interest) and sales expenses in period t
- BI_t : Business Inventories at the end of period t
- SD_t : Short Term Debt (incl. repayment of long-term debt) at the end of period t
- ALD_t : Additional Long term Debt obtained in period t

In his Cash Inefficiency model, to predict failure, he used 20 failed firms between 1954 – 1974 that were listed in Amsterdam Stock Exchange. 20 non-failed companies were paired

with similar asset size in similar industries. Financial data belonging to 5 years before failure were collected. In each of the 5 years, the contribution of the independent variables about liquidity were significant in predicting the failure.

Cash Flow Identity model was used by Aziz et al. (1988) .

The model can be stated as

$$(k_j - h_j) - (A_j + R_j - Y_j) - H_j - t_j = (F_j - N_j - M_j) + (D_j - B_j)$$

where

- $(k_j - h_j)$: Operating cash flow in year j
- $(A_j - R_j - Y_j)$: Net capital investment in year j
- H_j : Liquidity change in year j
- t_j : Taxes paid in year j
- $(F_j - N_j - M_j)$: Interest payments, long and short term debt raised or retired in year j
- D_j : Dividends paid in year j
- B_j : Equity capital raised in year j

So for the total market value of the company,

$$\sum_{j=1}^{\infty} \frac{(K_j - h_j) - (A_j + R_j - Y_j) - H_j - t_j}{(1+r)^j} = \sum_{j=1}^{\infty} \frac{(F_j - N_j - M_j)}{(1+r_d)^j} + \sum_{j=1}^{\infty} \frac{(D_j - B_j)}{(1+r_e)^j}$$

where

- r : weighted average cost of capital of the company
- r_d : cost of debt
- r_e : cost of equity

Mainly the left-hand side of the equation is the present value of expected entity cash flows after tax and right-hand side of the equation is the present value of expected lender cash flows + shareholder cash flows.

Entering into the model, all of the groups were divided by the book value of the firm for the scaling purposes.

49 failed companies that declared bankruptcy between 1971 – 1982 were selected again from the COMPUSTAT filings. These companies were matched with 49 non-bankrupt pairs without any restriction according to the size.

The data from one year to five years prior to the bankruptcy was extracted. From the descriptive analysis of the data for bankrupt and non-bankrupt group,

- All cash flows, except shareholders and liquidity change were higher for non-bankrupt firms in each of the five years prior to bankruptcy.
- Liquidity change cash flows were higher for non-bankrupt firms than for bankrupt firms in each of the five years except for year one prior to bankruptcy.
- Shareholder cash flows for bankrupt and non-bankrupt firms didn't not follow any pattern.

In the logit model with all of the variables, although the coefficients were obtained, unexpected directions (signs) of coefficients arose due to high multicollinearity between the variables. So, a second probit model, eliminating these multicollinearities were constructed for each year.

The negative sign of each independent variable in Table 2.6 below made sense that a potentially non-bankrupt firm was expected to generate higher operating cash flows, invest more in capital assets, pay out more in taxes, and be capable of carrying more debt than a matching bankrupt firm.

Table 2.6

Partial logit model coefficients

Years Prior to Bankruptcy	$k_j - h_j$	$A_j + R_j - Y_j$	t_j	H_j	$D_j - B_j$	$F_j - N_j - M_j$	Constant
1	-	-6.6805	-39.707	-	-	-125.93	2.5683
2	-	-7.7143	-	-	-5.2996	-17.375	0.71614
3	-5.2842	-	-27.739	-	-	-	1.6552
4	-	-9.6254	-40.595	-9.1608	-	-24.380	2.4263
5	-	-	-34.141	-19.230	-	-	1.3946

Source: (Aziz, Emanuel, & Lawson, 1988)

Table 2.7

Comparative classification accuracy between CFB model and ZETA and Z models

Years Prior to Bankruptcy	Z	ZETA	CFB	ZETA vs CFB	Z vs CFB
1	79.6	96.2	85.6	1.866	-0.797
2	75.5	84.9	79.6	0.702	-0.486
3	55.1	74.5	75.5	-0.117	<u>-2.121</u>
4	52.1	68.1	72.9	-0.528	<u>-2.105</u>
5	52.2	69.8	78.3	-0.959	<u>-2.629</u>

Significant at %1 level

Source: (Aziz, Emanuel, & Lawson, 1988)

The model was compared with Altman ZETA Model and Z Models in Table 2.7 above. In this comparison, it was seen that, the Cash Flow Model was outperforming the Z model in 5 years before bankruptcy (outperforming significantly in 3 years). The difference between ZETA and CFB model was not significant in none of 5 years, but CFB was slightly better in 3 years.

Trying to improve their previous results, Aziz and Lawson (1989) made another study, but this time adding the accounting basis financial ratios of Altman (1968). With 6 different hypotheses, they tried to test if their CFB model and their mixed model improved the overall classification, bankrupt company classification and result any bias.

- Hypothesis 1: CFB models do not perform better or worse than Z and Zeta models in the overall classification of bankrupt and non-bankrupt firms.
- Hypothesis 2: Mixed models do not perform better or worse than Z and Zeta models in the overall classification of bankrupt and non-bankrupt firms.
- Hypothesis 3: CFB models do not perform better or worse than Z and Zeta models in identifying potentially bankrupt firms.
- Hypothesis 4: Mixed models do not perform better or worse than Z and Zeta models in identifying potentially bankrupt firms.
- Hypothesis 5: CFB models do not misclassify potentially non-bankrupt firms at a different rate than the potentially bankrupt firms.
- Hypothesis 6: Mixed models do not misclassify potentially non-bankrupt firms at a different rate than the potentially bankrupt firms.

Constituting the estimation sample, 49 bankrupt companies from COMPUSTAT filings that declared bankruptcy between 1973 – 1982 were selected. They were matched with the 49

non-bankrupt companies according to industry and size criteria. 26 bankrupt and 67 non-bankrupt companies were also reserved as a holdout sample.

Their model was the Lawson Cash Flow Identity method again:

$$(k_j - h_j) - (A_j + R_j - Y_j) - H_j - t_j = (F_j - N_j - M_j) + (D_j - B_j)$$

where the cash flow variables were the same as the previous study as

- OPRG : $(k_j - h_j)$ Operating cash flow in year j
- CAPV : $(A_j - R_j - Y_j)$ Net capital investment in year j
- LIQH : H_j Liquidity change in year j
- TAXP : t_j Taxes paid in year j
- LNDC : $(F_j - N_j - M_j)$ Lender cash flows in year j
- SHRC : $D_j - B_j$ Shareholder cash flows in year j

In addition to the 6 CFB independent variables, Altman's Z Score model independent variables were also included into the model

- WCAT Net working capital to total assets
- REAT Retained earnings to total assets
- EBAT Earnings before interest and taxes to total assets
- MVEQ Market value of equity to book value of total debt
- SLAT Sales to total assets

Two Logit model constructions were made. In the first modelling, only the CFB variables were used, in the second model, all of the variables were used.

Table 2.8

Comparative Classification and Prediction Accuracy Between the CFB, Mixed, Z, and Zeta Models (in %)

Analysis Sample	Year	ZETA			Z			CFB			Mixed		
		Overall	Bankrupt	Nonbankrupt	Overall	Bankrupt	Nonbankrupt	Overall	Bankrupt	Nonbankrupt	Overall	Bankrupt	Nonbankrupt
Analysis Sample	1	92.8	96.2	89.7	88.8	79.6	93.9	91.8	85.7	97.9	95.9	93.9	98.0
	2	89.0	84.9	93.1	82.7	75.5	89.8	78.6	79.6	77.6	83.7	83.7	83.7
	3	83.5	74.5	91.4	74.5	55.1	93.9	77.6	75.5	79.6	80.6	81.6	79.6
	4	79.8	68.1	89.5	70.8	52.1	89.6	77.1	72.9	81.3	72.9	62.5	83.3
	5	76.8	69.8	82.1	67.4	52.2	83.7	71.9	78.3	65.1	70.8	71.7	69.8

Holdout Sample	Year	ZETA			Z			CFB			Mixed		
		Overall	Bankrupt	Nonbankrupt	Overall	Bankrupt	Nonbankrupt	Overall	Bankrupt	Nonbankrupt	Overall	Bankrupt	Nonbankrupt
Holdout Sample	1	77.4	80.8	76.1	76.3	88.5	71.6	76.3	88.5	71.6	82.8	92.3	79.1
	2	78.5	76.9	79.1	74.2	73.1	74.6	74.2	73.1	74.6	77.4	73.1	79.1
	3	71.7	53.9	78.8	76.1	80.8	74.2	76.1	80.8	74.2	81.5	88.5	78.8
	4	72.0	61.5	76.1	74.2	84.6	70.2	74.2	84.6	70.2	72.0	69.2	73.1
	5	70.6	62.5	73.8	71.8	83.3	67.2	71.8	83.3	67.2	72.9	79.2	70.5

Source: (Aziz & Lawson, 1989)

Analysing the estimation sample, with %95,9 overall classification rate, mixed model outperformed all other models in the prediction of bankruptcy one year prior to it. But in a general look, at every year prior to bankruptcy, ZETA model was the most superior one. For the Holdout sample, the mixed model had the highest accuracy in years one and three years prior respectively. the CFB model had the highest scores in years four and five respectively, and the Z model has the highest score in year two. Results are given in Table 2.8 above.

The Hypotheses were tested by the differences of the classification accuracies in Table 2.9 below.

Table 2.9

t Ratios to Compare the Accuracy of Different Bankruptcy Prediction Models

	Year	Analysis Sample				Holdout Sample	
		Zeta vs CFB	Zeta vs Mixed	Z vs CFB	Z vs Mixed	Z vs CFB	Z vs Mixed
Overall Accuracy Hypotheses 1 and 2	1	0.271	-0.961	-0.071	-1.870	0.178	-0.922
	2	2.053	1,119	0.727	-0.187	0.69	0.181
	3	1,079	0.546	-0.508	-1,023	-0.683	-1.579
	4	0.472	1,169	-0.994	-0.324	-0.338	0.000
	5	0.791	0.963	-0.653	-0.491	-0.181	-0.348
Bankruptcy Accuracy Hypotheses 3 and 4	1	1.866	0.537	-0.797	-2.088		
	2	0.702	0.167	-0.486	-1,007		
	3	-0.117	-0.864	-2.121	-2.820		
	4	-0.528	0.591	-2.105	-1,030		
	5	-0.959	-0.207	-2.629	-1.926		

Significant at %2 level
Significant at %5 level
Significant at %10 level

Source: (Aziz & Lawson, 1989)

Looking at the estimation sample, hypothesis 1 and 2 were not rejected, since very few of the classifications were significantly different. Holdout sample also didn't present significant differences between the models. So, in terms of overall accuracy, the ability to discriminate between bankrupt and non-bankrupt firms was about the same for all the models.

Hypothesis 3 and 4 were rejected because the CFB model and the mixed model presented significant differences when considered with Z Model, favouring the CFB and Mixed models. So it could be concluded that, the CFB models alone or mixed exhibited a superior predictive power in predicting the bankrupt companies, and offered a superior early warning capability. Hypothesis 5 and 6 offering that "the CFB and the Mixed models were biased toward bankruptcy" couldn't be rejected.

2.3.2 Prediction of Failure with Probabilistic Models

Some drawbacks and requirements of the statistical assumptions of Discriminant Analysis like linearity, normality and independence among variables caused the researchers to use probabilistic models like Logit and Probit. In fact, if these assumptions weren't violated, DA was more efficient than Probit and Logit (Lo, 1986). In Logit and Probit these assumptions were not needed to be validated. But these newer models were also criticised over a need to transform the original variables and the complicated computations involved.

Several drawbacks and associated problems of MDA models were mentioned by Eisenbeis (1977).

- There were certain statistical requirements on the distributions of predictors.
 - Variance and covariance matrices of the predictors should have been the same for both groups of bankrupt and non-bankrupt.
 - The requirement of normally distributed predictors prohibited the usage of the dummy variables.
- Since MDA was an ordinary discriminatory model, the output of it gave little intuitive interpretation.
- In MDA, pair sampling was mostly used, which brought several problems. The failed and non-failed firms were commonly matched according to the size and industry criteria. The gain and loss for applying such a sampling method or a different matching, or even no matching was not discussed in the literature. So, it might be better to include the sampling criteria into the model as a predictor.

Morrison (1969) populated a checklist of important considerations to examine, before making a Multiple Discriminant Analysis:

- Linear discriminant function was only appropriate when the groups' covariance matrices were equal or nearly equal.
- The Mahalanobis statistic tested only the statistical significance between groups. The statistic was heavily affected by the sample sizes.
- When the groups were unequal in size, it was difficult to interpret the classification tables, and as a consequence, in determining the samples, regardless of the total sample size, the effective sample size was governed by the smaller groups.
- In forming the classification decision (for example, in choosing the cut-off boundary values), the prior probabilities and opportunity costs of misclassification were to be considered.

Logit analysis solved most of the statistical problems associated with MDA. The estimation problem was reduced to the sample statement: given that a firm was belonging to some predefined population, what was the probability that the firm failed within some predefined time period? No assumptions had to be made about the distribution of predictors.

Ohlson's study (1980) was one of the pioneers of the probabilistic bankruptcy models. With the usage of logit model, he made several important contributions to the previous bankruptcy studies.

Ohlson commented that prior studies had carried out another major problem, besides the usage and assumptions of MDA. A realistic model required that the predictors were available for use before the event to be predicted. In case of the bankruptcy prediction, it was expected

that the independent variables were ready for use before the bankruptcy happened. But most of the time, since the annual reports were not publicly available at the end of the fiscal year (because the statements were to be audited), it was not clearly known that at what point in time the data were made available. Most of the prior search was using Moody's Manual which exhibited some uncertainty about the timing of the data. But he concluded that, for purposes of developing a discriminating methodology, super accurate data were not needed. The predictive quality of a model could be robust across a variety of data gathering and estimating processes.

But to overcome the risk mentioned above, Ohlson chose Chapter X or Chapter XI filer companies between 1970 – 1976 with the objective to obtain three years of data prior to the date of bankruptcy from the 10-K reports of the companies. In case the last available accountants' report explicitly stated that the company had filed for bankruptcy, a fourth report was also collected. His final sample was 105 bankrupt firms. In 18 of the companies, the last accountants' report was stating that the company had filed for bankruptcy, so this year records were deleted and previous 3 years' records were kept. So, the average lead time between the date of the fiscal year of the last relevant report and bankruptcy was longer than the previous studies and it was 13 months in average. To illustrate, Hers Apparel Inc. filed for bankruptcy at May 1974 and the last audited report ending the fiscal year at February 1974 was reported at July 1974, so the previous year's report dated April 1973 was chosen. The lead time between the last report and the bankruptcy was 13 months.

Pair sampling was not chosen. So the report year of the 2058 non-bankrupt companies were chosen randomly. 9 independent variables were chosen according to the popular usage in the previous literature.

- SIZE : log (Total Assets/GNP price level index)
- TLTA : Total Liabilities to Total Assets
- WCTA : Net Working Capital to Total Assets
- CLCA : Current Liabilities to Current Assets
- OENEG : 1 if Total Liabilities Exceed Total Assets, 0 otherwise
- NITA : Net Income to Total Assets
- FUTL : Funds provided by operations to Total Liabilities
- INTWO : 1 if net income was negative for 2 consecutive years, 0 otherwise.
- CHIN : $(NI_t - NI_{t-1}) / (|NI_t| + |NI_{t-1}|)$

$$l(\beta) = \sum_{i \in S_1} \log P(X_i, \beta) + \sum_{i \in S_2} \log (1 - P(X_i, \beta))$$

where

- S_1 is the set of bankrupt firms
- S_2 is the set of non-bankrupt firms
- $l(\beta)$ is the log likelihood of any specific outcomes
- $P(X_i, \beta)$ is the probability of bankruptcy
- $1 - P(X_i, \beta)$ is the probability of non-bankruptcy
- X_i is the vector of predictors
- β is vector of parameters

The aim was to find the β parameters where $l(\beta)$ is maximum

Three sets of models were calculated. (Table 2.10 below)

- Model 1: Predicting bankruptcy within one year
- Model 2: Predicting bankruptcy within two years
- Model 3: Predicting bankruptcy within one or two years

The significance of coefficients (except WCTA, CLCA and INTWO) were greater than 2 in all models so these coefficients were significant in predicting the bankruptcy.

Although slightly decreasing, the likelihood ratios of Model 1, 2 and 3 were 0,84, 0,80 and 0.72 respectively.

Table 2.10

Prediction Results

		SIZE	TLTA	WCTA	CLCA	NITA	FUTL	INTWO	OENEG	CHIN	CONST
MODEL1	Estimate	-0.407	6.03	-1.43	0.0757	-2.37	-1.83	0.285	-1.72	-0.521	-1.32
	t - statistic	-3.78	6.61	-1.89	0.761	-1.85	-2.36	0.812	-2.45	-2.21	-0.97
MODEL2	Estimate	-0.519	4.76	-1.71	-0.297	-2.74	-2.18	-0.78	-1.98	0.4218	1.84
	t - statistic	-5.34	5.46	-1.78	-0.733	-1.8	-2.73	-1.92	-2.42	2.1	1.38
MODEL3	Estimate	-0.478	5.29	-0.99	0.062	-4.62	-2.25	-0.521	-1.191	0.212	1.13
	t - statistic	-6.23	7.72	-1.74	0.738	-3.6	-3.42	-1.73	-3.11	1.3	1.15

Source: (Ohlson, 1980)

The significant result in all 3 models show that SIZE was an important predictor of bankruptcy. The negative sign of SIZE showed that, as the firm size got larger, the possibility of bankruptcy decreased.

The model significantly showed that the leverage (TLTA) increased the potential of a bankruptcy.

Although not significant in the short run, the negative sign of NITA showed that, the more profitable the companies were in the long run, the less probable they were to have a bankruptcy.

The study of Zavgren (1985) added several contributions for bankruptcy prediction models. Although she criticized some methodologies used in Ohlson's model, she built a dichotomous prediction model of logit like Ohlson's, predicting the bankruptcy 5 years prior to it. In addition, applying the entropy concept, Zavgren quantified the quantity of information contributed each year.

Conditional probability models derived the probability of a certain element belonging to certain dichotomous (or polytomous) classes, the coefficients showing the effect of a unit change in the independent variable on the probability of the dependent variable.

Comparative Advantages of Conditional Probability Models Over Discriminant Analysis were reported by her as

- If only a dichotomous prediction (like healthy or failing) was needed, this might require only the usage of a discriminant analysis model. But an investor might need more than a dichotomous result. As the investment opportunities were evaluated after a risk and return analysis, a probabilistic result would help more.
- From the perspective of a loan agreement, rather than a dichotomous result of a lend or no lend decision, the creditor would want to calculate the amount of a loan after making adjustments of risk. The likelihood of a failure only requires some adjustments on the risk premiums on the interest rate of the loan. Discriminant analysis could also generate a probability, but this usually involved subjective assessment of the probability associated with

a discriminant score. And even though the classification accuracy was high, the probabilities obtained from the discriminant function might be very inaccurate.

Zavgren employed the factors and related ratios to the factors which were found to have highest factor loadings in the Pinches et al. (1972) study, with one exception that, Zavgren preferred to use quick ratio instead of the current ratio (Table 2.11 below). (She said that the current ratio was a misleading measure because it increased as the failing firm's unsellable inventories accumulated, if quick ratio was used, a clearer picture would be obtained.)

Matching 45 failed and 45 non failed firms by their industries and asset sizes, the logit model was constructed and analysed on the perspectives of significance of the model, the significance of the probabilities, the significance of individual coefficients. The logit model was found to be significant as the likelihood ratio test statistic indicated significance at the 0.995 confidence level when compared with the Chi-square distribution with seven degrees of freedom.

Table 2.11

Factor Loadings of the Pinches et al. Study

Ratio	Factor	Factor Loading
Total Income / Total Capital	Return on Investment	0.97
Sales / Net Plant	Capital Turnover	0.95
Inventory / Sales	Inventory Turnover	0.97
Debt / Total Capital	Financial Leverage	0.99
Receivables / Inventory	Receivables Turnover	-0.99
Quick Assets / Quick Liabilities	Short Term Liquidity	0.81
Cash / Total Assets	Cash Position	0.91

Source: (Zavgren, 1985)

As for the individual coefficients, debt to invested capital was significant in all of the 5 years, and the positive sign was indicating that the debt usage had been increasing the probability of failure.

Turnover ratios had significant effect on bankruptcy in the long run. The positive signs of the coefficients were not because of the high levels of sales, but because of the lacking investment in the capital assets, and making the companies unable to compete with its competitors, and leading to an ultimate bankruptcy.

Zavgren applied Shannon's Entropy Theory to search for the informative ability of the model. Information Theory was introduced by Shannon to analyse the amount of information in the transmitted signals in an information system. Theil (1969) applied the information concept into economics by analysing the information in the financial statements.

“Entropy” is the degree of uncertainty over the occurrence of an event. Although the concept is mostly used in thermodynamics, by the introduction of the concept in information theory, it gained too much importance and area of application in research.

An entropy function is required to be a decreasing function of the probability of an event. A greater entropy in the probability distribution of an event will be taken to imply a greater degree of surprise if that event were to occur. It should also be additive to permit calculations of conditional measures. To calculate the expected entropy, each entropy can be weighted by its probability of occurrence.

$$H_n = H_n(p_1, p_2, p_3, \dots, p_n) = -\sum_{k=1}^n p_k \log p_k$$

Measuring the total entropies of the years before failure distinctly, Zavgren said that when the event of bankruptcy or non-bankruptcy was more certain (a smaller entropy value) or more surprising, uncertain (a bigger entropy value). For the failing firms,

- one year prior to failure, the total entropy was 20.05 and average entropy was 0.46 one year prior to failure,
- 5 years prior to failure, the total entropy was 24.56 and average entropy was 0.56 one year prior to failure,
- Meaning that in 5 year prior period, the failure was more uncertain (a greater entropy) and in one year prior, it was more certain.

After certain calculations, it was also an important finding that that the first derivative of the expected quantity of information was minus the logit index. In other words, the logit prediction was the negative of the rate of change in the quantity of information as the probability of the event changed.

Although they didn't seem to affect the statistical inferences or overall classification rates in bankruptcy prediction models, two bias caused by sample selection and data collection procedures employed in most financial distress studies were investigated and tried to be eliminated by Zmijewski (1984).

There were two estimation bias that could be observed in financial distress models:

- The “oversampling” of the distressed firms that happened because of the choice-based sample design. This issue arose because of the extremely low frequency of firms having financial distress. (Dun & Bradstreet Business Failure Record showed that the frequency rate of business failures in the US had never exceeded %0.75 since 1934.) A choice based sample

was observed, when the probability of an observation entering the sample had depended on the value, attribute or group of the dependent variable. In other words, a researcher first observed the dependent variable and then chose the sample based on that knowledge. This was done by the researchers because this method was cost effective when certain phenomena occur infrequently and when there was a high cost in collecting a random sample, containing a sufficient number of rare event observations. On the other hand this approach violated the random sampling design assumption. When a dependent variable group had a sample probability larger than the population probability (in other words, when the dependent variable group was oversampled), this group was expected to have understated classification and prediction error results.

- The “complete data sample” criterion arose because, the data for financially distressed firms were often unavailable. If financially distressed companies were less likely to have complete data, the probability of distress and the probability of having complete data should have negative correlation. The probability of distress given complete data would be less than the probability of distress observed in the population or in a randomly drawn sample.

17 studies about financial distress were examined (Table 2.12 below). All of the studies depended on the complete data criterion. And only 3 of these studies had financially distressed companies less than %40 of their total sample. So, it was clear that the models in these researches used non-random samples that had different compositions from the compositions of the population. If it was remembered that, the population frequency of a financial distress had never exceeded %0.75, then looking at the sample frequency rates, we could say that all of these studies had oversampled the distressed firm group (ranging from 2 times the population rate to 66.7 times the population rate). If a choice-based sampling bias was existing and oversampling a group would understate the error rates, we should expect

that, the Pearson Correlation Coefficient between the sampling frequency rate (of the distress group) and the classification error rates would be negative. The Pearson Correlation Coefficient of -0.932 significant at %1 level showed us that the group error rates were associated with sample frequency rates and pointing out a choice base sample bias.

3 independent variables were selected to predict bankruptcy

- ROA: Net Income to Total Assets (Return on Assets)
- FINL: Total Debt to Total Assets (Financial Leverage)
- LIQ: Current Assets to Current Liabilities (Liquidity)

Table 2.12

Sample Selection / Data Collection Procedures Used in Financial Distress Research

Research Study	Sample Selection	Sample Composition			Classification Error Rates		
	Complete Data Criterion	Distressed	Not-Distressed	Sample Frequency Rate	Distressed	Not-Distressed	Overall
Beaver [1966]	Yes	79	79	0.500	NR	NR	0.10
Altman [1968]	Yes	33	33	0.500	0.06	0.03	0.05
Wilcox [1971, 1973]	Yes	52	52	0.500	NR	NR	0.06
Deakin [1972]	Yes	32	32	0.500	0.03	0.03	0.03
Blum [1974]	Yes	115	115	0.500	NR	NR	0.08
Elam [1975]	Yes	48	48	0.500	NR	NR	NR
White and Turnbull [1975]	Yes	34	2303	0.015	0.53	0.01	0.02
Altman et al. [1977]	Yes	53	58	0.477	0.06	0.09	0.07
Deakin [1977]	Yes	63	80	0.441	0.12	0.01	0.06
Ketz [1978]	Yes	75	100	0.429	0.33	0.04	0.07
Van-Frederikslust [1978]	Yes	20	20	0.500	0.05	0.1	0.08
Norton and Smith [1979]	Yes	30	30	0.500	0.11	0.03	0.07
Dambolena and Khoury [1982]	Yes	46	46	0.500	0.09	0.01	0.06
Ohlson [1980]	Yes	105	2058	0.049	NR	NR	0.04
Emery and Cogger [1982]	Yes	52	52	0.500	NR	NR	0.10
Zavgren [1982]	Yes	45	45	0.500	NR	NR	0.18
Zmijewski [1983]	Yes	96	3880	0.024	0.83	0.01	0.02

Source: (Zmijewski, 1984)

There were several available estimation techniques appropriate on choice-based samples.

- Weighted Exogeneous Sample Maximum Likelihood (WESML)
- Conditional Maximum Likelihood (CML)
- Full Information Concentrated Maximum Likelihood (FICML)

WESML was the least complex of these techniques and also used in Zmijewski's (1984) study.

In WESML, the log-likelihood function was weighted by the ratio of the population frequency rate to the sample frequency rate of the individual groups.

For the Probit Assessment, the log likelihood function is

$$L = \sum_j (B) \ln[\Phi(H)] + \sum_j (1 - B) \ln[1 - \Phi(H)]$$

where

- B: 1 if bankrupt, 0 if otherwise
- Φ : Cumulative Density Function for Standard Normal Variable
- $H: a_0 + a_1ROA + a_2FINL + a_3LIQ$

In the WESML, the probit assessment is changed as

$$L = [POP / SAMP] \sum_j (B) \ln[\Phi(H)] + [(1-POP)/(1-SAMP)] \sum_j (1 - B) \ln[1 - \Phi(H)]$$

where

- POP: the proportion of bankrupt firms in the population.
- SAMP: the proportion of bankrupt firms in the sample.

In the WESML Probit Assessment, as the sample selection probability approached the population probability, the WESML likelihood function approached the unweighted likelihood function. Since studies with 50% sample frequency rate oversampled bankrupt group, So, $POP/SAMP \ll 1$ and $(1 - POP)/(1 - SAMP) \gg 1$ and this meant that the probability estimates of bankrupt group was upwardly biased and this group was overclassified.

The companies with SIC codes less than 6.000, that were listed on the American and New York Stock Exchanges between 1972 – 1978 were selected as the population of his study. A firm was included into the bankrupt group, if it had filed a petition during this period. The total number of bankrupt companies was 129, 81 of them had satisfied the complete data criterion. The bankrupt company frequency rates were distributed randomly between %0.49 to %0.94 within years. So, 81 bankrupt and 1.600 non-bankrupt companies with complete data were selected for the analysis.

From these data sets, two samples were separated. 40 bankrupt and 800 non-bankrupt as estimation sample, and 41 bankrupt and 800 non-bankrupt as prediction sample. 6 choice-based subsamples were created from the estimation sample. In these subsamples, 40 bankrupt companies were used but matched with 40, 100, 200, 400, 600 and 800 non-bankrupt companies respectively, to measure the effect of sampling frequencies in the groups. So the sampling frequency rates were %50, %28,6, %16,7, %9,1, %6,3 and %4,8 in each subsamples. The effects of a choice-based sampling were investigated using unweighted probit and WESML adjusted probit assessments by comparing the estimated parameters, the correct classification rates, and the correct prediction rates across estimation samples and techniques within these 6 sample groups (Table 2.13). If a bias didn't exist, then there should be no relation between the bankrupt firm sample frequency rate of the different estimation samples and the estimated parameters, classification and prediction error rates of the unweighted probit assessment.

Zmijewski's analysis had some very important implications. In the tables, Pearson correlation coefficients between the sample frequency rates and the coefficients and results in unweighted and WESML probit models are displayed.

In Table 2.14 below, the companies are classified as bankrupt or non-bankrupt considering a cut-off point of 0.5. These results also showed the existence of a bias so that

- The bankrupt firms were overclassified in the unweighted probit.
- In the unweighted probit, The correlation coefficient between the estimation sample frequency rate and the percentage of bankrupt firms classified correctly was 0.949. In the WESML, it was 0.797. This showed that WESML didn't eliminate the bias completely.
- The percent of companies, that were classified in the same group within unweighted and WESML assessment was also shown (Percent the same line) . This result showed that, classifications from the two techniques were more similar as the sample selection probability approached the population probability.

Table 2.13

Comparison of the Zmijewski Unweighted and WESML Probit Model Estimations Across Alternative Estimation Samples

	SAMPLES						Pearson Correlation Coefficient
	40:40	40:100	40:200	40:400	40:600	40:800	
	%50	%28,6	%16,7	%6,9	%6,3	%4,8	
Unweighted Probit							
C	<u>-2.953</u>	-3.674	<u>-4.206</u>	<u>-4.288</u>	<u>-5.503</u>	<u>-4.336</u>	0.859
ROA	<u>-9.023</u>	<u>-5.673</u>	<u>-6.669</u>	<u>-3.720</u>	<u>-3.918</u>	<u>-4.513</u>	<u>-0.908</u>
FINL	<u>5.503</u>	<u>5.606</u>	<u>6.083</u>	<u>6.126</u>	<u>5.917</u>	<u>5.679</u>	-0.650
LIQ	0.197	0.186	0.219	0.052	0.047	0.004	0.744
-2 x Log Likelihood Ratio	<u>76.89</u>	<u>107.74</u>	<u>147.35</u>	<u>176.24</u>	<u>190.87</u>	<u>203.78</u>	NA
WESML Probit							
C	<u>-4.777</u>	<u>-5.789</u>	<u>-5.922</u>	<u>-5.581</u>	<u>-5.106</u>	<u>-4.803</u>	0.143
ROA	<u>-8.702</u>	<u>-4.238</u>	<u>-5.019</u>	<u>-2.479</u>	<u>-2.739</u>	<u>-3.599</u>	<u>-0.915</u>
FINL	<u>5.133</u>	<u>6.236</u>	<u>6.574</u>	<u>6.496</u>	<u>5.725</u>	<u>5.406</u>	-0.361
LIQ	-0.002	0.146	0.213	0.031	-0.061	-0.1	0.21
-2 x Log Likelihood Ratio	<u>107.99</u>	<u>161.61</u>	<u>207.3</u>	<u>250.3</u>	<u>272.15</u>	<u>286.42</u>	NA

Significant at %1 level

Significant at %5 level

Significant at %10 level

Source: (Zmijewski, 1984)

- Since the statistical inferences of the overall model and individual coefficients were similar for both the unweighted and WESML assessments, it couldn't be concluded that the WESML approach improved the prediction results.
- In the unweighted assessment, all of the estimated parameters had statistically significant correlations with the sampling frequency. (Table 2.13 above)
- The 0.859 correlation of constant showed that the lower sample frequencies were associated with lower constant, pointing out a lower bankruptcy prediction probabilities. (Table 2.13 above)
- But in the WESML assessment, this high correlations between the sampling frequency and estimated parameters didn't exist. This showed that, the bias was existing in the unweighted probit but eliminated by the WESML.

As for the prediction sample, the classification results were like the estimation sample (Table 2.15 below). But there was a major difference that the non-bankrupt firm correlation with the sample frequency was -0.947, indicating a strong under classification bias. The non-bankrupt companies were under classified as the sampling frequency of non-bankrupt was decreasing. This indicated that, when parameters which were less biased were used, more correct overall classification rates were to be reached.

Table 2.14

Comparison of the Zmijewski Unweighted and WESML Probit Model Classifications Across Alternative Estimation Samples

	SAMPLES						Pearson Correlation Coefficient
	40:40	40:100	40:200	40:400	40:600	40:800	
	%50	%28,6	%16,7	%6,9	%6,3	%4,8	
Unweighted Probit							
Bankrupt	92.50	77.50	77.50	72.50	65.00	62.50	<u>0.949</u>
Nonbankrupt	100.00	100.00	99.00	99.30	99.50	99.50	0.681
Overall	96.30	93.60	95.40	96.80	97.30	97.70	NA
WESML Probit							
Bankrupt	52.50	47.50	50.00	48.00	42.50	42.50	0.797
Nonbankrupt	100.00	100.00	99.50	100.00	99.80	99.90	0.310
Overall	76.30	85.00	91.30	95.00	96.30	97.10	NA
Percent the Same	80.00	91.40	95.00	96.80	98.30	98.70	-0.989

Significant at %1 level

Significant at %5 level

Significant at %10 level

Source: (Zmijewski, 1984)

Table 2.15

Comparison of the Zmijewski Unweighted and WESML Probit Model Classifications in the Prediction Sample Across Alternative Estimation Samples

	SAMPLES						Pearson Correlation Coefficient
	40:40	40:100	40:200	40:400	40:600	40:800	
	%50	%28,6	%16,7	%6,9	%6,3	%4,8	
Unweighted Probit							
Bankrupt	97.60	90.20	90.20	78.00	70.70	70.70	<u>0.895</u>
Nonbankrupt	92.50	97.90	98.60	99.00	99.10	99.50	<u>-0.947</u>
Overall	92.60	97.60	98.30	98.10	97.90	98.20	<u>-0.892</u>
WESML Probit							
Bankrupt	46.30	48.80	51.20	46.30	41.50	43.90	0.538
Nonbankrupt	99.80	99.60	99.60	99.60	100.00	100.00	-0.281
Overall	97.30	97.10	97.40	97.10	97.30	97.40	-0.129

Significant at %1 level

Significant at %5 level

Significant at %10 level

Source: (Zmijewski, 1984)

Complete Data Criterion: The Complete Data Criterion was examined by estimating the model using both a simple probit assessment (ignoring the incomplete data problem) and a bivariate probit assessment (incorporating the effects of incomplete data into the estimation). The bivariate probit model consisted of two probit equations, the financial distress equation and the complete data equation. As the result, it was shown that, firms with high bankruptcy probabilities had more tendency to have incomplete data. So again, a bias was clearly shown to exist, but, in general, it didn't appear to affect the statistical inferences or overall classification rates

2.3.3 Hazard Model and Prediction of Failure with Time Series Data

It is clear that the models in predicting bankruptcy commonly use the financial data at an instant of time. The models try to predict bankruptcy certain amounts of time before it happens.

As mentioned by Shumway, in this cross-sectional data selection, the fact that firms change through time was ignored. Shumway (2001) developed a Hazard Model, that used all available information to determine the bankruptcy risk of each company at every instant of time. The dependent variable of the hazard model was the time spent by the company in the healthy group. Hazard model was built to incorporate time-varying explanatory variables. The model produced more efficient out-of-sample forecast by utilizing more data. The difficulty of estimating the model because of the nonlinear likelihood functions and time-varying covariates was the disadvantage of the model.

To estimate the hazard model, each year the firm survived (was not bankrupt) was included into the sample data as a firm that didn't fail. Each bankrupt company contributed to the model as bankrupt only once.

From the analysis with the hazard model, Shumway concluded that, half of the accounting ratios used in Altman and Zmijewski models were not good predictors. On the contrary, the previously neglected market-driven variables like market size, past stock returns, standard deviation of the stock returns were good estimators of failure.

Sun (2007) also used Hazard Model in his study. Although he outperformed the models constituted with auditor opinions with his model, as a limitation for hazard models, he underlined that the significance of the predictor variables across time were due to change.

Kahya and Theodossiu (1999) also criticized the prior models about their handling of the cross-sectional data. Although it was known that the financial data used in the prediction studies were expected to possess serial correlation over time, their values at only an instance of time were used as independent variables in the models. They argued that the business failure prediction could be made more accurately with the models that account for

- The serial correlation in the data
- Information from more than one period
- The stationary explanatory variables.

For this purpose, they used Cumulative Sums (CUSUM) model on 117 healthy and 72 failed companies between 1974 – 1991. They concluded that, in a time-series manner, growth of deflated total assets, change in the ratio of fixed assets to total assets, and change in the ratio of operating income to sales had positive relationship and change in the ratio of inventory to sales had a negative relationship with the financial performance.

2.3.4 Evaluations and Re-tests of the Known Models

The bankruptcy prediction models that was pioneered by Beaver with univariate use of ratios, Altman with the usage of MDA model, and evolved by the use of Probabilistic Regression Models (pioneered by Ohlson), were frequently re-tested and re-estimated with newer data by academicians.

Deakin (1972) re-tested Beaver's Dichotomous Classification and MDA. He applied Dichotomous Classification Test on 32 failed and 32 non-failed companies between 1964 – 1970 and obtained the same results. There were high correlations between the classification errors obtained by Beaver and Deakin for the years prior to bankruptcy.

He also applied MDA for the 14 ratios used in Beaver's study. The analysis showed that, the MDA model showed discriminative ability with high accuracies in 3 years prior to failure.

Begley et al. (1996) re-tested the original models of Altman and Ohlson, and re-estimated the models with newer data to obtain new coefficients. These new coefficients had important implications of underlining the changing usage of leverage and increased importance of liquidity.

They used 1980 data to examine whether newer data resulted different classification errors when known mentioned models were re-evaluated. Their sample were 165 bankrupt and 2.000 nonbankrupt companies of 1980 – 1989 period.

With the original model re-tested with 1980 data, Type I and Type II errors both increased in Altman's model. As for Ohlson's model, similar Type I error rate to the original study, but an increased Type II error rate was observed.

Then the models were re-estimated with newer coefficients. The coefficient of leverage variable changed in a way to have a smaller effect on the probability of failure. That was because, there was an increasing acceptance of relatively high corporate debt levels and the perceived benefits of additional leverage increased in 1980s when past was considered. And

bankruptcy laws were used strategically by companies. Though the coefficients of the leverage variables were showing a predictive ability, these strategic uses of bankruptcy were resulting more classification errors and less emphasis on the related coefficients.

Second important change was more emphasized role of the liquidity variable, consistent with the growing focus on cash flows in 1980's.

For the prediction of failure in Turkish Textile Sector Altaş and Giray (2005) tried to develop their logistic model. 5 ratios of liquidity, 11 ratios of financial position, 9 ratios of performance and 8 ratios of profitability were selected for 2001. Factor Analysis were applied for the dimension reduction and the new dimensions were introduced as the independent variables into the logistic model. It was concluded that, for the textile sector in 2001, the most important factor to determine the failure was the liquidity factor.

2.3.5 Advanced Techniques in Prediction of Failure using Artificial Intelligence and Machine Learning

Increasing computational power and recent trends in the development of “Data Mining”, “Machine Learning” and “Artificial Intelligence” had 2 important consequences:

- New alternatives developed in solving nonlinear problems employed in prediction.
- New class of researchers involved in the topic. Although the area of bankruptcy prediction was mostly attracting the researchers in the finance and statistics areas, after the mentioned development in technology, it had also been a major area of interest for the researchers of ML, AI, Data Analytics which were mostly the profession of Computer and

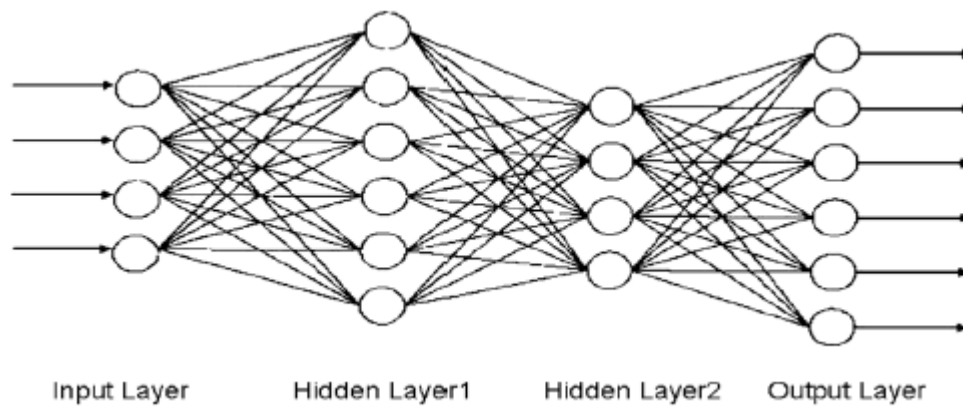
Data Engineers. This was an expectation for future research mentioned by Bellovary et al. (2007). The improvements of the technology broadened the class of researchers of the topic. The intelligent models like expert systems, fuzzy logic, decision trees, genetic algorithm (GA), support vector machine (SVM) and neural networks were used in decision making aiding the managers.

Chen, Huang, & Lin (2009) offered that, fuzzy logic could describe a problem using more human-like reasoning methods and so tolerate inaccuracy and uncertainty, but employed some difficulties in obtaining the correct knowledge base. Neural networks could learn from observed data but the output was a black box and causal relationship between the variables were not demonstrated, so the outputs couldn't be easily modelled and used by management. Neural networks had disadvantages like their bad prediction performance in the out-of-sample data while their performance in in-sample data were good.

So, they proposed a hybrid, neuro fuzzy system, to construct a bankruptcy early warning system, which combined the functionality of fuzzy logic and the learning ability of neural network.

There were many different topologies of neural networks. Back-propagation was the most commonly used topology. It drew the mapping function between the input and output from the provided data set. The mapping function was complicated to explain the causal relationship between the variables. There were 3 types of layers displayed in Figure 2.2: input layer, hidden layer, and output later. Every node in the input layer represented an independent variable. A node in the output layer represented the dependent variable. The nodes in the hidden layers existed to complete the nonlinear transformation calculation.

Figure 2.2

Layers of a Neural Network

When the obtained model was employed in forecasting, each node in the input layer sent its value to the hidden layer. Each node in the hidden layer was expected to calculate the weighted sum of the input values according to the post-training edge weight, and perform the nonlinear transformation of the sigmoid (logistic) function to produce an output, which was the input of the next hidden layer. Then each node in the second hidden layer and each node in the output layer would repeat the same procedure to produce an output. The final output of the node in the output layer would be the final output of the model.

The purpose of back-propagation was to obtain the weight of each edge, in order to minimize the squared error sum between the actual value and the predicted value. First each edge was given a random value. Then, the squared error sum between the actual value and the predicted value could be calculated. Next the weight was updated according to the gradient search method until the squared error sum was less than or equal to the threshold value. So, the eventual model was trained, and could be used to forecast.

100 bankrupt companies filing for Chapter 11 in the period 1998–2002 were included in the sample. The banking industry was excluded from their sample. Their financial data were obtained from the most recent seasons prior to the date of filing. Matched sampling was used and 100 healthy firms according to their size, industry and time period were also chosen. The total sample of 200 failed and non-failed companies were divided into 2 subcategories of 140 companies (70 bankrupt, 70 healthy) and 60 companies (30 bankrupt, 30 healthy) training sample and the test sample respectively.

16 financial ratios categorized into six dimensions were used as independent variables: Short-term liquidity ratio (STLR), cash flow ratio (CFR), capital structure and long-term solvency ratio (CSLTSR), return on investment ratio (ROIR), assets utilization ratio (AUR) and operation performance ratio (OPR) to measure an enterprise's operation.

First, factor analysis was applied to reduce the number of variables used in the model. Each variable was standardized by subtracting the average divided by their standard deviation before conducting the factor analysis. The Factor Analysis showed that all of the ratios were representative so all of the input factors were retained. (Table 2.16)

Table 2.16

Summary of Factor Analysis

Factor	Name of the Factor	Variables	Definition of Variables	Factor Loadings	Eigen Values	Cumulative % of Variance
F1	STLR	X1	Current assets/current liability (CA/CL)	0.961	1.848	92.39%
		X2	(Cash + cash equivalents)/current liability ((C + CE)/CL)	0.961		
F2	CFR	X3	Cash flow/total debts (CF/TD)	0.875	2.209	55.22%
		X4	Retained earnings/total assets (RE/TA)	0.845		
		X5	(Current assets- current liabilities)/total assets ((CA-LA)/TA)	0.648		
		X6	Cash flow/total assets (CF/TA)	0.555		
F3	CSLTSR	X7	Total debts/total assets (TD/TA)	0.898	2.271	56.77%
		X8	Total debts/total equity (TD/TE)	0.87		
		X9	EBIT/(EBIT – interest expense (EBIT/(EBIT-IE))	0.761		
		X10	Operating income/interest expense (OI/IE)	-0.358		
F4	ROIR	X11	Income before taxes/average total assets (IBT/AOA)	0.809	1.308	65.40%
		X12	Operating income/interest expense (OI/IE)	0.809		
F5	AUR	X13	Net sales/total Assets (NS/TA)	-	-	-
F6	OPR	X14	(Sales – cost of sales)/sales ((S-CS)/S)	-	-	-

Source: (Chen, Huang, & Lin, 2009)

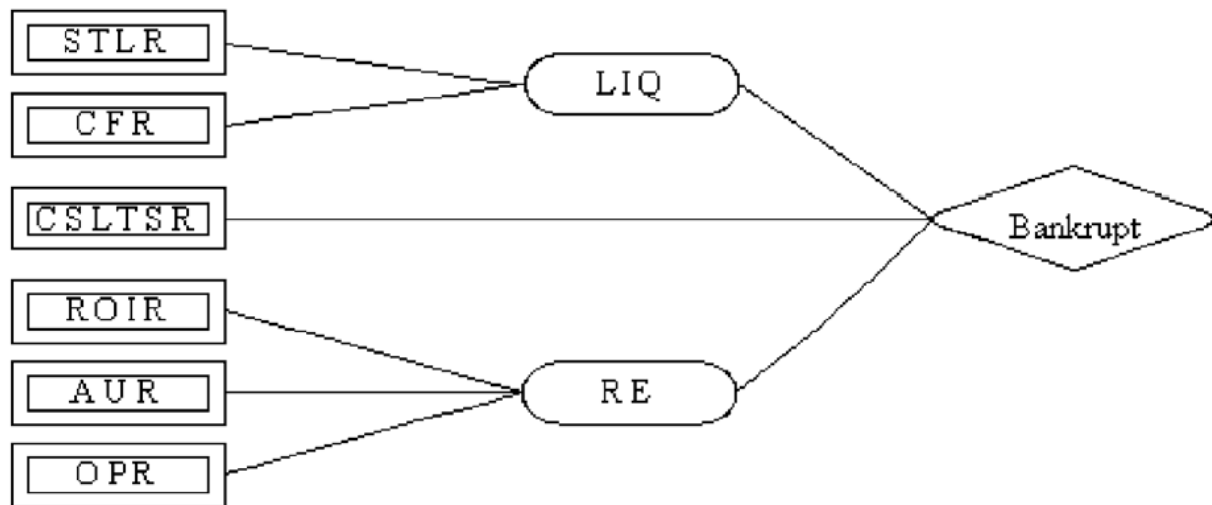
Second, the neuro fuzzy model was constructed as combination of factors of

- STLR and CFR leading to a factor of LIQ (the liquidity)
- ROIR, AUR, and OPR produce a factor of RE (the retained earnings)
- CSLTSR (capital structure and long-term solvency)

From the logit analysis, neither CFR, CSLTSR, AUR, nor OPR variables were found to be significantly associated with bankruptcy. But a more complicated logit model incorporating interaction effects between the same variables were constructed. This new model showed that, firms with lower returns relative to investment and higher risk to liquidity were more likely to go bankrupt, but also bankruptcy prediction was sensitive to the interactions between variables of cash flow ratio and short-term liquidity ratio, return on investment ratio and operating performance ratio.

Figure 2.3

Combination of Factors in the Neuro Fuzzy Model



Type I and II errors were defined as in opposite structure compared to previous study. Type I error was defined as considering a healthy company as bankrupt one, and Type II Error as considering a bankrupt company as healthy). From these definitions, as the cost of a Type II error was greater than that for a Type I error, a sensitivity analysis was performed by changing a multiple of 50,100, 200, 500 and 1.000 between Type II / Type I costs, to investigate if the cost gap between a Type I error and a Type II error affected the model performance.

$$\pi = (k_1B + k_2C) = (k_1B + mk_1C) = k_1 (B + mC)$$

where

- k_1 : The misclassification cost of a Type I error
- k_2 : The misclassification cost of a Type II error
- π : Total expected misclassification cost

Neural Network and Neuro Fuzzy classifications were more accurate than logit in both training and test samples. In addition, when the cost of misclassification was considered, Neural Network and Neuro Fuzzy algorithms caused less cost in misclassification when compared with logit both in training and test samples again. (Tables 2.17 – 2.18)

Table 2.17

The empirical results for training data

Quarters	Accuracy Rate			Misclassification Cost			Detecting Power		
	Logit (%)	NN (%)	NF (%)	Logit	NN	NF	Logit (%)	NN (%)	NF (%)
Q1	76.43	82.86	78.57	1,022	221	722	85.71	97.14	90.00
Q2	77.14	80.00	78.57	1,022	423	1,019	85.71	94.29	85.71
Q3	74.29	74.29	82.86	1,228	431	42	82.86	94.29	100.00
Q4	68.57	76.43	77.86	835	626	429	88.57	91.43	94.29
Q5	66.43	70.71	75.71	1,139	634	46	84.29	91.43	100.00
Q6	65.71	72.14	68.57	1,444	830	937	80.00	88.57	87.14

Source: (Chen, Huang, & Lin, 2009)

Table 2.18

The empirical results for testing data

Quarters	Accuracy Rate			Misclassification Cost			Detecting Power		
	Logit (%)	NN (%)	NF (%)	Logit	NN	NF	Logit (%)	NN (%)	NF (%)
Q1	75.00	80.00	81.67	312	309	308	90.00	93.33	90.00
Q2	76.67	71.67	83.33	512	215	209	83.33	93.33	93.33
Q3	71.67	66.67	66.67	716	515	319	76.67	83.33	90.00
Q4	65.00	68.33	68.33	615	514	415	80.00	83.33	86.67
Q5	66.67	61.67	68.33	420	320	225	86.67	90.00	96.67
Q6	58.33	65.00	61.67	519	516	412	83.33	83.33	86.67

Source: (Chen, Huang, & Lin, 2009)

With the improvement of computational power, The Decision Tree has been another analytic method, used in bankruptcy prediction in recent years. There were several different algorithms followed under Decision Tree Methodology. In Chen's research (2011), 3 of

Table 2.19

The accuracies of Chen Model with different level factors across different algorithms

	C5.0				CART				CHAID				LR			
	Prior Period	Nonbankruptcy	Bankruptcy	All	Prior Period	Nonbankruptcy	Bankruptcy	All	Prior Period	Nonbankruptcy	Bankruptcy	All	Prior Period	Nonbankruptcy	Bankruptcy	All
No Factor Analysis	2	94.44%	100%	97.01%	2	91.66%	100%	95.83%	2	92.06%	93%	92.29%	2	86.11%	84%	85.07%
	4	90.47%	92.53%	91.50%	4	90.47%	94.02%	92.24%	4	87.30%	92.53%	89.91%	4	88.88%	89.55%	89.21%
	6	91.95%	86.17%	88.95%	6	79.31%	77.65%	78.45%	6	73.56%	86.17%	80.11%	6	85.05%	94.68%	90.06%
	8	85.58%	91.53%	88.80%	8	81.08%	90.00%	85.89%	8	89.18%	87.69%	88.38%	8	91.89%	91.53%	91.70%
1st level factor analysis	2	95.45%	90%	92.53%	2	96.77%	92%	94.03%	2	91.66%	94%	92.54%	2	86.11%	87%	86.60%
	4	90.47%	91.04%	92.53%	4	88.88%	91.04%	91.04%	4	87.30%	89.55%	88.42%	4	88.88%	89.55%	89.21%
	6	88.50%	80.85%	84.53%	6	78.16%	72.34%	75.14%	6	73.56%	78.72%	76.24%	6	88.50%	81.91%	85.08%
	8	81.08%	83.07%	82.07%	8	84.68%	84.61%	84.65%	8	82.88%	86.15%	84.65%	8	86.48%	83.84%	85.06%
2nd level factor analysis	2	88.88%	97%	92.54%	2	91.66%	97%	94.03%	2	91.66%	93%	92.09%	2	75.00%	87%	80.60%
	4	88.65%	92.53%	90.59%	4	92.06%	89.55%	90.77%	4	87.30%	91.04%	89.23%	4	82.53%	82.08%	82.31%
	6	89.65%	79.78%	84.53%	6	78.16%	71.27%	74.59%	6	85.05%	86.17%	85.64%	6	87.35%	78.72%	82.87%
	8	81.08%	82.08%	81.58%	8	89.18%	83.07%	85.89%	8	81.08%	87.69%	84.65%	8	85.58%	77.69%	81.74%

Source: (Chen M.-Y. , 2011)

these algorithms were compared with logistic regression model in Table 2.19. These 3 algorithms were

- C5.0
- Classification and Regression Trees (CART)
- Chi-squared Automatic Interactive Detector (CHAID)

Mainly, his objective in the study was

- Developing Decision Tree and Logistic Regression models to predict financial distress
- Using non-financial independent variables (as well as financial ratios) to increase prediction accuracy
- Using Principle Component Analysis (PCA) and analyse if reduction in the variables had any effect on prediction accuracy

The sample was 100 companies (50 financial distress, 50 non-financial distress companies) listed in Taiwan Stock Exchange Companies (TSEC) List, between January 2000 to May 2007. 37 variables were used that were categorized in 6 main groups as:

- Earning ability: Pre-tax Margin, Return on Total Assets, Return on Equity, Earnings per Share, and Gross Margin.
- Financial structure ability: Debt to Assets, Times Interest Earned, Book Value Per Share, Financial Leverage ratio, Debt to Equity, Short Term & Long Term Debt to Book Value ratio, Fixed Assets to Total Assets, Gross Margin to Total assets, Inventory to Total assets, Inventory to Sales, Investment ratio, and Current Assets to Total Assets.
- Management efficiency ability: Inventory Turnover, Account Receivable Turnover, Fixed Assets Turnover, Total Assets Turnover, Equity Turnover, Working Capital Turnover.

- Management performance: Pre-tax Margin Growth, Gross Margin Growth, and Sales Growth.
- Debt-repaying ability: Current Ratio, Acid test Ratio, Cash Ratio, Cash Flow Ratio, Cash Flow to Long Term Debt, Cash Flow to Total Debt, and Cash Flow to Short Term and Long Term Debt.
- Non-financial factors: Dividend Pay out Ratio, Price to Book Value Ratio, the proportion of collateralized shares by the board of directors, and the Insider Holding Ratio.

As mentioned by Son et al. (2019), there were several reasons that the number of independent variables were tried to be reduced.

- More independent variables brought more computational difficulties and burdens to be faced like incompleteness and missing data.
- As the number of independent variables increased, it got harder to interpret the results.

On the other hand, when the model performance was considered, reducing the number of variables was not desirable because information loss existed while removing some variables.

To reduce dimensionality, PCA was used. In the first run, 21 values were eliminated due to low communality. With the remaining 16 variables, 84.4% of total variance was explained. One more level of factor analysis was run, reducing the independent variables to 12 and the total variance explained was %94. The remaining 12 variables were Return on Equity (ROE), Earnings per Share (EPS), Return on Asset (ROA), Cash Flow Ratio, Cash Flow to Total Debt Ratio, Current Ratio, Acid-Test Ratio, Inventory to Total Assets Ratio, Current Assets to Total Assets, Gearing Ratio, Debt to Equity Ratio, and Debt/Equity (DE).

Then each of these independent variable sets were introduced to Decision Tree Models and Logistic Regression Model separately. The comparisons of models in Table 2.19 showed that

- There was a deterioration in the in the Decision Tree Models' precision accuracy as the years prior to financial distress got larger.
- There was no deterioration in the Logistic Regression Model.
- Logistic Regression Model was better than the Decision Tree Models in the long run precision, although the Decision Tree was more successful in the short run.
- In the Decision Tree Models, as the factors were decreased and less independent variables entered into the models, the precision deteriorated.
- Since the error in the prediction of a failing company was more costly than error in prediction of a non-failing company, Decision Tree models were more successful because they possessed less errors in predicting the bankrupt ones.

Son, Hyun, Phan and Hwang (2019) used Korean Companies listed in Korean Credit Rating Agency NICE Information Service between 2011 and 2016 in their research. There were 977.940 companies, of which 23.137 were bankrupt.

2 categories of data were used:

i.The basic information about company

- Company basic principal code
- Size classification code
- Listed Stock Market
- External Financial Auditing

ii.108 financial information about the companies

- Growth ratios

- Leverage ratios
- Liquidity ratios
- Operational ratios
- Profitability ratios
- Solvency ratios
- The other variables. The variables that do not belong to any of the seven categories in the above

Five algorithms were used and compared

- Logistic Regression (LR)
- Random Forest (RF)
- Gradient Boosting (XGBoost)
- Gradient Boosting (LightGBM)
- Artificial Neural Network (ANN).

The presence or absence of an external audit was a meaningful classification dimension because it represented various attributes, such as the size of a company. So, for each algorithm, 3 datasets were calculated separately

- the full dataset
- data for companies that had been externally audited
- data for companies that had not been externally audited.

After the definitions of

- TP: True Positive. The model correctly identified a firm that went bankrupt
- FN: False Negative. The model failed to identify a firm that went bankrupt

- FP: False Positive. The model incorrectly classified a firm as having gone bankrupt, when it did not
- TN: True Negative. The model correctly identified a firm that did not go bankrupt.

Performance of the algorithms were compared by the help of ROC curves . A ROC (Receiver Operator Characteristic) Curve is a graphical plot that illustrates the diagnostic ability of a binary classifier.

ROC curve was generated by plotting TPR against FPR where

$$\text{True Positive Rate (TPR)} = \frac{TP}{TP+FN}$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP+TN}$$

The area under the curve (AUC) was calculated for each model and for the validation set, a higher value of AUC meaning a better accuracy.

It can be seen from the comparison in Table 2.20 that

- All of the new models (Random Forest, XGBoost, LightGBM, Artificial Neural Network) achieved higher accuracy than the Logistic Regression Model. All of these models improved accuracy measure by AUC up to %25
- In addition, the analysis showed that, the Korean companies with external audits had some attributes (such as size, etc) that were helpful in prediction of bankruptcy.

Table 2.20

The Comparison of the Different Algorithms

MODEL	DATASET	AUC	IMPROVEMENT
Logistic Regression	ALL	0.67	
	+EX	0.78	
	NOEX	0.67	
Logistic Regression (*)	ALL	0.75	12%
	+EX	0.85	9%
	NOEX	0.74	10%
Random Forest	ALL	0.77	15%
	+EX	0.88	13%
	NOEX	0.75	12%
XGBoost	ALL	0.83	24%
	+EX	0.88	13%
	NOEX	0.82	22%
LightGBM	ALL	0.83	24%
	+EX	0.87	12%
	NOEX	0.83	24%
Artificial Neural Network	ALL	0.82	22%
	+EX	0.85	9%
	NOEX	0.81	25%

(*) The same variables were introduced a Box-Cox transformation, to eliminate the skewness in the data, and Logistic Regression was rerun. The analysis showed that this transformation increased the accuracy of the data when considering the case of entering into the Logistic Regression without this transformation.

Source: (Son, Hyun, Phan, & Hwang, 2019)

Koh and Low (2004) compared Logistic regression, Neural Network and Decision Tree Algorithms in accurately predicting the going concern predictions of auditors. 165 non-going concern and 165 going concern evaluations were analysed. In the training sample, the accuracies of three different models were very close for logistic regression, neural network and decision tree as %95,22, %95,65 and %97,39 respectively. But for the validation sample, the accuracies were %94.00, %91.00 and %95.00 respectively. So, the decision tree algorithm was superior in both evaluation and validation samples.

2.4 Costs Related with Bankruptcy

Modigliani and Miller (1958) stated in their famous Capital Structure Irrelevance Theory that, the value of a firm was irrelevant to the capital structure choices made by the management of a company. Rather than the leverage, the market value of a firm was dependent on the operating profits of the company. But the theorem leaned on several assumptions that were violated in real conditions like

- No taxes
- No transaction costs
- No bankruptcy costs
- No asymmetric information between investors.

So this means that, in real life, bankruptcy costs and probability of a bankruptcy was a phenomenon that was to affect the capital structure decision. In fact, the optimal capital structure decision was a trade-off between the tax subsidies arising from the deductibility of interest and the bankruptcy costs. (Kalaba, Langetieg, Rasakhoo, & Weinstein, 1984)

Haugen and Senbet (1978) discriminated between bankruptcy and liquidation of a firm. Bankruptcy was the formal transfer of authority about companies' management from the investors to the creditors and the reorganization of the capital structure due to the unmet obligations to the creditors. During this transfer, direct and indirect costs might arise. Direct costs were mainly the legal, accounting and trustee fees. Indirect costs were the opportunity costs resulting from the disruption of firm's relations between its suppliers and customers. Independent from bankruptcy, liquidation was a capital budgeting decision (made by whoever was in control), the sale of the dismantled assets whose market value exceeded their aggregated value as a continuing firm.

Leaving the Modigliani and Miller assumption of no bankruptcy costs, Warner stated that when a bankruptcy prospect appeared, these costs might affect the optimal capital structure decision. But he mainly investigated the direct costs related to bankruptcy. When the court finalised the decision of any type of bankruptcy, the claimholders were expected to hire agents to maximize the values of their respective claims. These agents included lawyers, accountants and various professional consultants. The direct costs of bankruptcy were mainly attributed to these costs.

To investigate these costs, Warner (1977) analysed 11 bankrupt railroad companies, which were reported by ICC. The costs included payment to all parties for legal fees, professional services, trustee fees and filing fees. In these analysis, it was seen that on average, the bankruptcy case was settled in 13 years on average, and the average direct cost was \$1.88 million. There were other several implications of his analysis introduced in Table 2.21.

The data shows that the mean debt/equity ratio increase sharply because as the company declared bankruptcy, the market value of equity fell very rapidly whereas market value of debt didn't.

Table 2.21

Change of the Market Value and Bankruptcy Costs (in Railroad Companies) as Bankruptcy Approaches

	Months before Bankruptcy				
	0	-12	-36	-60	-84
Average Debt/Equity Ratio	24.3	18.2	13.7	7.8	3.8
Market Value (million\$)	50.0	87.1	156.0	216.7	251.3
Bankruptcy Cost/Market Value	%5.3	%4.0	%2.5	%1.4	%1.0

Source: (Warner, 1977)

Although the bankruptcy costs were getting substantial as the bankruptcy date was near, these costs were not directly proportional to the market value. The percentage declined for relatively high market values. The R^2 between the market value of the company and bankruptcy costs / market value was 0.88. This evidence suggested that there were substantial fixed costs associated with the bankruptcy process.

Edward Altman is famously known with the development of his Z and ZETA Models, that quantitatively aim to predict the corporate failure. But in developing these models to predict failure, he also made empirical (rather than theoretical) assessments about costs that could arise when a bankruptcy happened.

To assess the costs of bankruptcy, Altman (1984) investigated the direct costs (costs including legal, accounting, filing and other administrative costs arising with bankruptcy) and the indirect costs (mainly the opportunity costs of the lost profits that a firm could be expected to suffer due to significant bankruptcy potential). Joining these costs with the probability of bankruptcy, the expected present value of cost of a bankruptcy was obtained. This expected cost of bankruptcy was compared with the expected present value of tax benefits from leverage. This comparison had important implications for the decision of whether or not an optimum capital structure was achieved. If bankruptcy costs were significant then it might be argued that at some point, the expected value of these costs outweighed the tax benefit gained from increasing leverage and the firm would have reached its optimum capital structure point.

2.4.1 Direct Costs of Bankruptcy

Direct costs were the explicit costs paid by the debtor in the reorganization/ liquidation process. These costs were the costs of legal, accounting, filing and other administrative costs arising with bankruptcy. These costs were documented in the bankruptcy records of individual firms but, no aggregate records were kept and it was necessary to go through the files in the Bankruptcy Courts, etc. In his study, this had been done for a sample of 12 retailer and 7 industrial firms.

2.4.2 Indirect Costs of Bankruptcy

Indirect Costs related to bankruptcy were not mentioned in literature, due to the opportunity cost nature and the difficulty in specifying and empirically measuring such costs.

An aspect of indirect costs was the lost sales and profits resulting when potential buyers of a product or service perceive that default was likely. Altman emphasized this phenomenon with real cases from several US Companies (like Chrysler, Franklin National Bank receivership, and the Braniff International failure cases): The President of Chrysler said that the share of new car sales dropped nearly two percentage points just because potential buyers feared the company would go bankrupt. In addition, company management spent a great deal of time and expense to store up the public's confidence in Chrysler as a continuing entity.

Indirect costs occurred even after the firm has declared bankruptcy and also when it was attempting to manage a return to financial health after bankruptcy. In addition, these costs were not limited to firms which actually did fail. Firms with high probabilities of bankruptcy, whether they eventually fail or not, still incurred these costs due to customer wariness, dealer problems, etc.

Suppliers might be reluctant to continue to sell to the high-risk companies without significant restrictions on their contracts and high prices. Some managerial cost increases like the costs of negotiations, security precautions and monitoring of debt covenants, and other debt-related agency costs that were incurred to protect the priority of creditors, might offset the tax advantage of debt and led to a choice of an optimal capital structure.

From the investor point of view, assuming that the markets were perfect, information was equally available to potential investors and all securities were available for all investors, a controversy showed up that, owners were able to diversify their investments and reap the gains from a safer competitor firms' gains when the potential customers of the endangered company bought from them. (For example, the potential buyers of Chrysler bought from Ford,

when Chrysler in endangered) . So a perfect substitution of goods and services were expected and that the loss from one firm was fully compensated by the gain from others. Also, prices would reflect these shifts such that investors would be equally as well off. Costs to society should be zero despite the bankruptcy possibility. But, markets were not perfect and investment opportunities were not always available in competitor firms (especially when the market was as internationally competitive one as automobiles, tractors, and airlines, and alternative investments in foreign firms were not easily transacted, or available).

Indirect Costs of Bankruptcy were calculated by 2 methods in his study.

- Regression Procedure
- Security Analyst Forecasts.

2.4.2.1 Indirect Costs Calculation with Regression Analysis

Calculation of indirect costs of bankruptcy with “Regression” employed several steps

- a. Each bankrupt company’s sales were used as the dependent variable, and appropriate industry sales were used as the independent variable, and regressed for the 10-year period prior to the bankruptcy.

$$S_{i,t} = a + b S_{I,t} \quad \text{for } t=10 \text{ years}$$

where

- $S_{i,t}$: Sales of the firm i in period t
- $S_{I,t}$: Aggregate sales of the industry I in period t

For example if we were to estimate the indirect costs for

- the third year prior to failure, sales of the firm were regressed on industry sales for the period t-13 to t-4,
- the second year prior to failure, sales of the firm were regressed on industry sales for the period t-12 to t-3,

For the Retail Sector, average R^2 of the regression of the industry sales over the 12-bankrupt company sample was 0,78 and for the industrial Companies sample of 7 companies, it was 0,46.

- b. From the average historical Profit Margin on the same 10 year period and the estimated sales for t-3, t-2, t-1 and t, the estimated profits for the estimated sales amounts were calculated.

$$\hat{S}_{i,t} = a + b S_{I,t} \quad \text{for } t=-3, -2, -1, 0$$

$$\hat{P}_{i,t} = \hat{S}_{i,t} PM$$

where

- $\hat{S}_{i,t}$: Estimated Sales
- $\hat{P}_{i,t}$: Estimated Profits
- PM: Average Historical Profit Margin

- c. From the estimated profits and the observed profits, unexpected profits (or losses) were calculated to reach the indirect costs of bankruptcy.

$$\Delta P_{i,t} = P_{i,t} - \hat{P}_{i,t}$$

where

- $\Delta P_{i,t}$: Unexpected Profits (losses)
- $P_{i,t}$: Observed Profits

- $\hat{P}_{i,t}$: Estimated Profits
- d. The total value of the firm was measured by adding the market value of equity to the market value of debt plus the book value of other debt plus the capitalized value of financial leases.
- e. The indirect and the direct costs of bankruptcy were compared to the total value of the firm for up to three years prior to failure.

Table 2.22

Direct and Indirect Costs of Bankruptcy vs the Company Value

	(Indirect + Direct Costs) / Value				Direct Costs/ Value				Indirect Costs/ Value			
	t-3	t-2	t-1	t	t-3	t-2	t-1	t	t-3	t-2	t-1	t
Retailing Companies	8.70%	8.20%	8.70%	12.20%	3.10%	3.00%	2.70%	4%	5.60%	5.20%	6.00%	8.20%
Industrial Companies	17.40%	17.20%	15%	23.70%	6.20%	7.30%	7.50%	9.80%	11.20%	9.90%	7.50%	13.90%

Source: (Altman, 1984)

In the conclusion, it was clear that, as the bankruptcy approached, the costs of bankruptcy were at a nontrivial level of nearly %25 of the company value. (Table 2.22 above)

2.4.2.2 Indirect Cost Calculation as measured by Security Analysts

The indirect bankruptcy costs estimates were calculated by the unexpected earnings (losses) for the period up to three financial statements prior to bankruptcy.

Indirect Costs = Analyst Earning Estimates – Actual Earnings.

Based on this calculation, the Indirect Costs of Bankruptcy reached %20 – 30 of the value of the company. But it was stated that there was a problem for the significance of the method because, as a firm approached bankruptcy, the number of analysts following the company tended to diminish.

When these 2 methods were considered, the total bankruptcy costs were not less than the %25 or %30 of the value of the company.

2.4.2.3 Expected Bankruptcy Costs vs Tax Benefits Trade off

In the decision of if the company reached the optimal capital structure or not, Comparison of the expected present value of bankruptcy costs to the expected present value of tax benefits derived from debt capital was to give the answer.

$$\frac{P_{B,t} BC_t PV_t}{MV_t} \quad \text{vs} \quad \frac{T_c iD_t PV_t (1 - P_{B,t})}{MV_t}$$

where

- $P_{B,t}$: Probability of a bankruptcy estimated in period t
- BC_t : Direct and Indirect costs of bankruptcy estimated in time t
- T_c : Marginal tax bracket of the company
- iD_t : Interest Expense from period t to infinity
- PV_t : Present Value adjustment back to period t
- MV_t : Market Value of the company in time t

The comparison of expected present values of bankruptcy costs vs. tax benefits from leverage is listed in the Table 2.23 and 2.24 below. The probabilities of bankruptcy were calculated

with Altman Zeta Model. The interest rate and levels of debt was assumed to remain constant from the measured time to infinity. In Table 2.23 and 2.24 below, the expected present value of tax benefits was calculated separately both on interest bearing debt and on the combined total of interest-bearing debt and capitalized leases. These tables show the comparison 2 and 1 year prior to bankruptcy respectively. If bankruptcy costs exceeded the tax benefits, then it was concluded that the company had too much leverage in its capital structure and its optimum point. These firms were overleveraged and in their optimal capital structure, an important component to be considered additionally was the bankruptcy costs.

In their ZETA Model, Altman et al. (1977) made major contributions to previous studies about bankruptcy prediction. They introduced a new bankruptcy model because of several new needs.

- The size of the financial profile of business failures had changed since Altman's 1968 study. The average size of bankrupt firms had increased with the greater visibility and concern from financial institutions, regulatory agencies and the public. Most of the past studies used smaller firms in their samples.
- Past failure models concentrated on manufacturers or specific industries. In their study, they also included the retailing companies on an equal basis with manufacturers.
- The data and footnotes to financial statements were prepared to adopt the most recent changes in financial reporting standards and accepted accounting principles.
- Testing and assessment of the Multiple Discriminant Analysis after several developments was needed, that constructed a base for past bankruptcy studies but was still very controversial.

Table 2.23

Comparison of the Tax Benefits from PV of Bankruptcy Costs vs. PV of Interest on Debt and Capitalized Leases (At t-2 Prior to Bankruptcy)

Company	Probability of Bankruptcy	Expected Present Value of Bankruptcy Costs(\$000) (3)	Expected Present Value of Tax Benefit from Debt (\$000) (4)	Expected Present Value of Tax Benefit from Debt & Capitalized Leased (\$000) (5)	3/4	3/5
1	0.77	1,749	559	1,328	3.13	1.32
2	0.63	1,618	1,769	32,120	0.91	0.05
3	0.61	1,628	673	1,073	2.42	1.52
4	0.34	519	959	959	0.54	0.54
5	0.32	547	468	8,241	1.17	0.07
6	0.27	262	1,577	1,577	0.17	0.17
7	0.89	19,179	3,104	11,495	6.18	1.67
8	0.9	6,279	1,406	1,406	4.47	4.47
9	0.27	2,075	10,722	14,137	0.19	0.15
10	0.63	1,360	428	3,865	3.18	0.35
11	0.33	609	7,240	30,780	0.08	0.02
12	0.48	206	95	182	2.17	1.13
13	0.48	7,680	50,400	88,730	0.15	0.09
14	0.97	818	84	179	9.74	4.57

Source: (Altman, 1984)

Table 2.24

Comparison of the Tax Benefits from PV of Bankruptcy Costs vs. PV of Interest on Debt and Capitalized Leases (At t-1 Prior to Bankruptcy)

Company	Probability of Bankruptcy	Expected Present Value of Bankruptcy Costs(\$000) (3)	Expected Present Value of Tax Benefit from Debt (\$000) (4)	Expected Present Value of Tax Benefit from Debt & Capitalized Leased (\$000) (5)	3/4	3/5
1	0.78	1,673	88	1,345	19.01	1.24
2	0.95	2,200	271	1,867	8.12	1.18
3	0.99	2,782	75	97	37.09	28.68
4	0.79	1,111	439	439	2.53	2.53
5	0.65	1,023	240	5,919	4.26	0.17
6	0.57	530	978	978	0.54	0.54
7	0.99	19,257	366	889	52.61	21.66
8	0.99	7,400	140	140	52.86	52.86
9	0.98	8,180	427	521	19.16	15.70
10	0.95	2,014	398	959	5.06	2.10
11	0.54	1,029	3,888	19,720	0.26	0.05
12	0.89	581	12	30	48.42	19.37
13	0.73	11,910	7,450	28,940	1.60	0.41
14	0.99	549	37	70	14.84	7.84

Source: (Altman, 1984)

New accounting practices they applied during this study were

- Capitalization of leases
- Adjustment for reserves
- Netting of minority interest and other liabilities on the balance sheet.
- The consolidation of captive finance companies and other non-consolidated subsidiaries.
- Deduction of goodwill and intangibles
- Writing of research and development costs, capitalized interest as expense

Their sample of bankruptcy were 53 companies of manufacturing and retail industries of 1969 – 1975. The first contribution was that their ZETA model classified bankrupt firms more successful (Table 2.25). The correct classification rates were %90 and %70 respectively for one year and 5 years prior to bankruptcy. The statistical method used was Multiple Discriminant Analysis (MDA). Linear and quadratic discriminant methods were tested separately. After introduction of 28 variables, 7 representative variables were selected as

- Return On Assets: $\text{EBIT} / \text{Total Assets}$
- Stability of Earnings: Standard Error Estimate of $(\text{EBIT} / \text{Total Assets})$
- Debt Service: $\text{Log} (\text{EBIT} / \text{Total Interest Payments})$
- Cumulative Profitability: $\text{Retained Earnings} / \text{Total Assets}$
- Liquidity: Current Ratio
- Capitalization: $\text{Market Value of Equity} / \text{Total Capital}$
- Size: $\text{Log} (\text{Assets})$

After ranking the performance of these variables in 6 tests (forward stepwise discriminant analysis, backward stepwise discriminant analysis, scaled vector test, separation of means

test, conditional deletion test and univariate F-test), the most important and discriminative variable was selected as Cumulative Profitability Ratio, followed by Stability of Earnings, and Capitalization variables. Although still contributing to the discriminative ability, Return on Assets was the least important.

There were some important implications of the study:

- ZETA Model had the predictive ability up to 5 years prior to bankruptcy. (%70 accuracy in predicting the bankrupt companies 5 years prior to their bankruptcy.)
- Classification accuracy of the ZETA model was higher than the previous (Altman, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, 1968) model.
- Linear model was equally accurate with the quadratic model in the estimation sample, but had superior accuracy in the holdout sample.

Table 2.25

Comparison of the Classification Accuracies between ZETA and Z Models

Years Prior to Bankruptcy	Zeta Model		1968 Model	
	Bankrupt	Non-Bankrupt	Bankrupt	Non-Bankrupt
1	% 96.2	% 89.7	% 93.9	% 97.0
2	% 84.9	% 93.1	% 71.9	% 93.9
3	% 74.5	% 91.4	% 48.3	n.a
4	% 68.1	% 89.5	% 28.6	n.a
5	% 69.8	% 82.1	% 36.0	n.a

Source: (Altman, Haldeman, & Narayanan, 1977)

Another important implication of this study (which also constitutes a baseline for this thesis paper) was its attempt to find the expected cost related with the estimation of ZETA and the optimal cut-off point to minimize this cost.

The optimal cut-off point for ZETA was assumed to be

$$\text{ZETA} = \ln \left(\frac{q_1 c_I}{q_2 c_{II}} \right)$$

where

- q_1, q_2 : Prior probabilities of bankruptcy and non-bankruptcy respectively
- C_I, C_{II} : Costs of Type I and Type II Errors respectively.

When equal prior probabilities of q_1, q_2 and equal amounts of costs for C_I, C_{II} were assumed (that is, $q_1 / q_2 = 1$ and $C_I / C_{II} = 1$), the optimal cut-off point for ZETA was to be equal to 0.

But in real, q_1 is not equal to q_2 and C_I is not equal to C_{II} .

$$EC_{\text{ZETA}} = q_1 (M_{12} / N_1) C_I + q_2 (M_{21} / N_2) C_{II}$$

where

- M_{12}, M_{21} : Observed Type I and Type II Errors respectively
- N_1, N_2 Number of observations in in Bankrupt and non-bankrupt groups respectively

In any given year, the ratio of the number of failed companies to the non-failed was % 0.5. So q_1 / q_2 was expected to be between 0.01 – 0.05 range.

In addition, the definition of Type I Error region cost was made as “the cost of an accepted loan that defaults”, Type II Error region cost was made as “the cost of a rejected loan that would have a successful payoff”. So, the empirical base to constitute these definitions were specified as

$$C_I = 1 - \frac{LLR}{GLR}$$

$$C_{II} = r - i$$

where

- LLR: amount of loan losses recovered
- GLL: gross loan losses
- r: effective interest rate on the loan
- i: opportunity cost for the bank

These definitions were important because these were the first theoretical definitions for the costs of Type I and Type II regions in the literature.

At ultimate, the commercial banks took the risk of losing all or a portion of a loan when the applicant would default. So, C_I was obtained from a research, covering annual report data of 26 largest US Commercial Banks. Encompassing the 5 year period of 1971 – 1975, C_I was assumed to be %70. (Where the applicant eventually default, only %30 of the amount of the loans were recovered by the banks.)

Since C_{II} was mostly an opportunity cost, the loss might be mitigated by another alternative use in other loanable funds. The banks' opportunity cost implied another cost at the same risk that would probably payoff. So, in this respect C_{II} was extremely small. It was assumed to be $r - i > 0$, but very small.

So the C_{II} was assumed to be %2, and C_I / C_{II} to be 35 times.

With these new findings about q_1 / q_2 and C_I / C_{II} ,

$$\text{ZETA} = \ln \left(\frac{q_1 c_I}{q_2 c_{II}} \right) = \ln \left(\frac{0.02 \times 0.70}{0.98 \times 0.02} \right)$$

$$\text{ZETA} = -0.337$$

It was sure that this optimal cut-off ZETA point would change for different values of q_1 / q_2 and C_I / C_{II} ratios

Although other banking services had been developed, commercial and industrial lending process was still the lifeblood for commercial banks. The journey of the lending process, from the initial loan request to the successful or unsuccessful repayment of the loan was investigated by Altman (1980).

Using the data of 5.000 banks for period between 1969 to 1976, reflecting their charge-off and recovery records, and an analysis of 400 individual charged-off loans belonging to 49 banks, his study was aiming to

- investigate the association between the credit evaluation stage and the following events that describes and quantifies the charge-off and (if existed) recovery experience.
- take on account the difference of costs in Type I and Type II Error regions and specify the cost of lending errors for bank loans and the optimal cut-off score for credit scoring.

The landing process consisted of the stages

- Application for the loan
- Credit evaluation
- Loan review
- Repayment performance

If the repayment failed, the processes of charge-off and workout followed.

The aim of the credit evaluation was to determine repayment probabilities, to provide the basis for the pricing and structuring of the loan agreement by using qualitative methods, ratio and industry analysis. Another way of thinking for the credit evaluation was the firm insolvency analysis, in which the likelihood of the continuation of the existence of the company was tried to be measured.

When these methods were used, mostly the historical records and experience on similar loans were used by the banks, but the cost of making errors were not considered.

Most common quantitative model for assessing bankruptcy prediction and risks of loan repayment was Multiple Discriminant Analysis. In MDA, the cut-off point for the dichotomous bankrupt or non-bankrupt decision (differently saying, lend or no lend decision) was defined as

$$CS = \ln \left(\frac{q_1 c_I}{q_2 c_{II}} \right)$$

where

- CS: Optimal Cut-off Point
- q_1, q_2 : Prior probabilities of bankruptcy and non-bankruptcy (or an observation would repay or not repay) respectively
- C_I, C_{II} : Costs of Type I and Type II Errors respectively.

Additionally

$$C_I = 1 - \frac{LLR}{COA}$$

where

- LLR: amount of loan losses recovered
- COA: Charged off amount

So C_I might be defined as the proportion of the charged-off loan that was not recovered.

Although precise calculation of these costs was very difficult, it was known that in the following specification of C_I , the result will be understated. Because, the opportunity costs of the bank on the other forgone interest revenues, the bank officers' time spent on lending operations and legal charges and expenses were neglected.

To quantify C_I , 3 different analyses were made.

1. A questionnaire about the costs of lending was applied to regional banks and 25 large banks who returned their results between 1971 to 1975. The result was that, before the subsequent recoveries of the charged-off loans, average loss was %77 for regional banks, and %83 for large banks. And after the recoveries, the loss was %69 for regional banks, and %72 for large banks. On average the rate of losses (C_I) was specified as %70.
2. The previous analysis was based on small numbers of banks. In the further analysis, working with New York Federal Reserve data, the recovery rates on loan losses for period 1969 – 1976 from the entire population of member banks (over 5,000 banks) were available in Table 2.26.

For some banks, the loss ratio was extremely high. For this reason, the average ratio of C_I was high relative to the median values.

The average of the median values of the 8 years was %29.0. So, the loss ratio was calculated to be %71 (which was very compliant with the previous analysis)

Another implication of the study with Federal Reserve Data was, the relationship between the bank size and the loan loss recovery experience (Figure 2.4). Larger banks recovery rates were less on charged-off loans than smaller ones. Larger banks maintained the unpaid amounts longer while others write off at an earlier stage. That might be because that larger banks give larger-sized loans which were a part of a lending consortium, so it might require more time for the bank to take care of the unpaid loan.

3. In the third analysis the recovery was broadened by the inclusion of bank's efforts to collect prior to the actual charge off. For instance, in the prior analysis, if a loan of \$100 was lent and the bank collects \$20 prior to charge-off, the charged-off amount would be \$80. But if another \$20 was recovered after the charge-off, the mentioned methods would show %25 recovery rate, and %75 loss rate. But in real, \$40 of the lent amount (\$100) was recovered, showing a recovery rate of %40 (or a loss rate of %60). The methodology in third calculation took this approach into consideration.

$$\mathbf{LLR} = 1 - \frac{\alpha + \beta}{LB_t}$$

where

- LLR: Rate of loss on loans,
- α : Amount collected prior to charge-off,
- β : Amount recovered after charge-off, and
- LB_t : Loan Balance at some point t prior to charge-off.

Table 2.26

Recovery Rates on Loan Losses

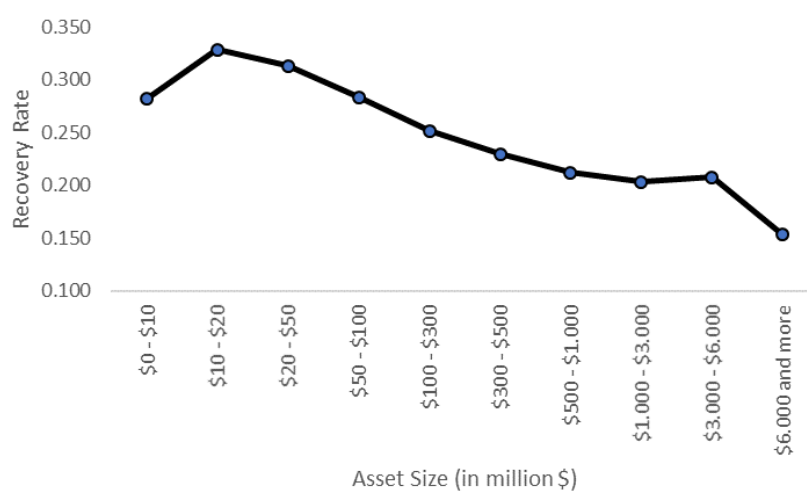
	1969	1970	1971	1972	1973	1974	1975	1976	8-Yr Avg
Number of Banks	4,792	4,893	4,889	4,950	5,022	5,160	5,128	5,252	5,022
Average Recovery Ratio	0.522	0.457	0.543	0.696	0.598	0.492	0.549	0.588	0.556
Median Recovery Ratio	0.304	0.265	0.317	0.341	0.304	0.255	0.267	0.283	0.290

Source: (Altman, 1980)

In our example,

$$\text{LLR} = 1 - \frac{20\$ + 20\$}{100\$} = 60\%$$

Figure 2.4

Median Recovery Rates by Bank Asset Size, 1969 – 1976

For the analysis, a short questionnaire was sent to bank representatives in the New York, South eastern, and Northern California regions of the United States. Over 400 questionnaires were sent and 55 replied back, of which 49 were usable. In the questionnaire, each bank was asked to select up to three charged-off loans for 1975, 1976 and 1977 for each and to provide

- the loan balance 12 months prior to charge-off,
- the balance charged-off,
- the amount recovered after charge-off.

For the 377 loan records of the banks

- (1) $\frac{LB_{12} - COA}{LB_{12}}$
- (2) Recovered amount after Charge Off/ LB_{12} ,
- (3) Total Recovery Rate (1+2)
- (4) Traditional Recovery Rate = Recovered Amount after Charge Off/Charge Off Amount

were calculated

LB_{12} : Loan Balance amount at the 12 months prior to charge off

COA: Charged off amount

The results (Table 2.27) below showed that the recovery performance of banks' bad loans charge-off was on average %37.6. Or, the loss rate (Type I error on defaulted loans was around % 62) In the previous findings, this rate was found to be around %70. So, in this research, the banks recovery performance seemed to be slightly better.

Table 2.27

Collection and Recovery Experience on Sample of Charged-off Loans (1975 – 1977)

	$\frac{LB_{12} - COA}{LB_{12}}$	Recovery / LB_{12}	(1) + (2)	Recovery / COA
	(1)	(2)		
Mean ¹	%16.5	%21.1	%37.6	%25.7
Mean ²	%38.4	%12.3	%50.7	%20.0

Mean¹ representing the average of the calculated ratios, Mean² representing the calculation based on the accounts consisting the formulations. So in Mean², larger-sized loans were weighted heavily.

Source: (Altman, 1980)

Returning back to $CS = \ln \left(\frac{q_1 C_I}{q_2 C_{II}} \right)$, if we assumed the prior probability of bankruptcy q_1 as %1.5 and cost of Type II error as %2, the optimal cut off would be

$$CS = \ln \left(\frac{0.015 \times 0.62}{0.985 \times 0.02} \right) = -0.75.$$

As C_I was lower than the previous studies, the optimal cut-off score was also lower, indicating a more liberal crediting policy.

Using a Probit model, Koh (1992) investigated the misclassification costs in bankruptcy prediction models and if these misclassification costs required to change the cut-off point used in the probit model or not.

There were some statistical steps in bankruptcy prediction models that included several problems:

- The application of discriminant analysis
- Usage of choice-based sample

- The selection of independent variables

The remedies for these common problems in the same order were

- The Probit analysis
- Weighted exogeneous sample maximum likelihood (WESML) or other adjustment procedures to eliminate the effects of the choice-based sampling
- Theoretical background of bankruptcy prediction models

In most of the models, the aim was to maximize the number of correct predictions, or in other words, to minimize the number of companies in Type I (the misclassification of a non-going concern as a going concern) and Type II (the misclassification of a going concern as a non-going concern) regions in total. But there were different misclassification costs in these error regions. And there was a strong tendency on reducing the Type II errors at the expense of the Type I errors which were much more costly.

There were different misclassification costs, because, when a non-going concern was predicted as a going concern (When a Type I error was made) the investor might lose the entire investment. On the contrary, when a going concern was predicted as a non-going concern (when a Type II error was made) the investor mostly lost an opportunity, the dividend income or a capital gain. Most of the times, even this cost was not expected to occur, because this investor would probably substitute the investment forgone with another investment of the risk level. This meant that, the misclassification cost of a Type I error was much higher than the misclassification cost of a Type II error. So the cut-off point that minimized the total number of Type I and Type II error regions in total was not optimal. At the extreme, if we

think that the cost of a Type I error was infinity, the optimal cut-off point should be where there were no Type I error occurring.

In Koh's research, a logit model was developed to predict the going concern, with a matched sample. From the COMPUSTAT records, 165 nonfinancial companies that filed for bankruptcy between 1980 – 1985 were paired with 165 companies with a going concern on the basis of industry, size and fiscal year.

6 independent variables were used:

- Quick assets to Current Liabilities (QACL)
- Market Value of Equity to Total Assets (MVTA)
- Total Liabilities to Total Assets (TLTA)
- Interest Payments to Earnings Before Interest and Tax (IEBT)
- Net Income to Total Assets (NITA)
- Retained Earnings to Total Assets (RETA)

So the theoretical index was

$$Z = 1.7791 + 2.0677 \times QACL + 0.9424 \times MVTA - 9.2441 \times TLTA - 0.0070 \times IEBT + 18.420 \times NITA + 2.2530 \times RETA$$

and

$$Y = F(Z)$$

where

- F: Cumulative logistic probability function

Table 2.28

Results of the Koh Logit Model

Variable	Coefficient	t-value	Probability >t
Constant	1.7791	2.4656	0.009
QACL	2.0677	3.5304	0.001
MVTA	0.9424	2.9910	0.003
TLTA	-9.2431	-4.3033	0.001
IEBT	-0.0070	0.0056	0.996
NITA	18.4200	4.7636	0.001
RETA	2.2530	2.2222	0.027

Except IEBT, all of the independent variables were significant at %5 level.

Source: (Koh H. C., 1992)

Second, he developed the Expected Misclassification Cost model as

$$EC = PN \times PI \times CI + PG \times PII \times CII$$

where

- EC : Expected misclassification cost of using the model
- PN : Prior probability of non-going concerns
- PG : Prior probability of going concerns
- PI : Conditional probability of Type I errors = (number of Type I errors given a particular cut-off point)/(number of non-going concerns) = (number of Type I errors)/165
- PII : Conditional probability of Type II errors = (number of Type II errors)/165
- CI : Misclassification cost of a Type I error

- CII : Misclassification cost of a Type II error

Here, PN and PG were calculated as 0.006 and 0.994 according to the rate of the companies filing for bankruptcy between 1980 – 1985 in COMPUSTAT records.

CI and CII were not measured. Because, the costs of predicting a going concern were generally intangible and very hard to measure (It was measured in the (Altman, Haldeman, & Narayanan, 1977) and (Altman, 1980)) . Instead, relative ratios for CI and CII were assigned ranging from 1:1 to 500:1 (Meaning that, the cost of a misclassification in Type I error region changed 1 time to 500 times the cost of an error in Type II error region)

From the final analysis, it was concluded that, although the optimal cut-off points were affected and changed by different misclassification costs, these cut-off points that minimized the expected misclassification costs were not sensitive to every change in the misclassification costs (Table 2.29) .

Table 2.29

Summary of Optimal Cut-off Points

From CI: CII	To CI: CII	Z
1:1	71:1	-4.0197
72:1	139:1	-3.2822
140:1	191:1	-2.6311
192:1	251:1	-1.4663
252:1	500:1	-1.1001

Source: (Koh H. C., 1992)

2.5 Legal Definitions of Bankruptcy

The legal framework of bankruptcy was explained in detail by Rüzgaresen (2011).

2.5.1 Turkish Legal Perspective

In the Turkish Law, the event of “Bankruptcy” can be explained and categorized from different points of view as

2.5.1.1 According to the way the bankruptcy is declared

- **Pursued Bankruptcy:** Pursued Bankruptcy can be started when the unpaid debt continues even the legal bankruptcy payment order is sent to the debtor by the creditor. The previous step for a creditor to start a pursued bankruptcy is starting a legal announcement of a bankruptcy proceeding.
- **Direct Bankruptcy:** Direct bankruptcy can be followed by the bankrupt party or the creditor without announcement of a bankruptcy proceeding in several cases. But in the bankruptcy code, the formal, and usual process is the pursued bankruptcy, whereas the direct bankruptcy is an exception.

2.5.1.2 According to the Laws and Codes about bankruptcy

- Enforcement and Bankruptcy Code
- Turkish Commercial Code
- Civil Law
- Banking Law

2.5.1.3 According to the initiator of the bankruptcy process

- Debtor
- Creditor
- Both parties

2.5.1.4 According to the debt is due or not

- As mentioned, as a general rule, the unpaid due supplied by the creditor is the main starting point of a bankruptcy process
- But in some cases, the bankruptcy process can be started without waiting for an undue payment, followed by a bankruptcy proceeding. When the location of the debtor is indefinite, when the debtor has the suspicion of defaulting its commitments and liabilities, when the debtor is defaulting the bankruptcy process can be initiated.

2.5.2 American Legal Perspective

The ultimate aim of American Law about bankruptcy is to ensure the creditors to get their full receivable, and to help the debtor continue its commercial presence at the same time. In the American Law, Bankruptcy is one of the results of the bankruptcy process initiated.

Bankruptcy is defined under 2 distinct processes for the companies. (There are other chapters related with bankruptcy mostly involving the personal or municipal entities.) These are:

- Chapter 11 Bankruptcy: This type of a bankruptcy process doesn't necessarily involve the closure of a business. But rather, the organization of the company and its assets are reorganized so that the company can be able to pay its debts to its creditors.
- Chapter 7 Bankruptcy: This type of the bankruptcy involves the "Liquidation" process. That is the selling of the company assets and property to pay its debts to the creditor.

The main difference of American Bankruptcy Code is that its main focus on reorganization of company and continuity of its legal entity. In American Law, the bankruptcy result is the exception.

But although the main focus is reorganization, such a focus may have anomalous consequences. The competition in the long-run should cause inefficient firms to be eliminated. So as the failing firms are more likely to be inefficient than firms in general, it would be rewarding the inefficiency and offsetting the beneficial effects of competition to aid the survival of failing firms.

So, to mitigate these effects, a new bankruptcy code went into effect by 1979. White showed the effects of this new code (1983). He argued that with the new Code, reorganization remained still as an option for failing firms, but made more difficult to use, increasing economic efficiency, and decreasing the aggregate bankruptcy costs. New code restricted the power of managers to continue economically inefficient enterprises. Managers and equity holders had a tendency toward reorganization rather than the economic efficiency considerations because they lost their jobs and/or their shares. As stated by them, the percent of liabilities of failed firms involved in reorganizations fell from 48% to 35% by the new code. The total aggregate bankruptcy costs per year had a decrease from \$35.6 billion to \$23.2 billion

3. METHODOLOGY

3.1 Selection of the Companies

Companies listed in Borsa Istanbul between 1998 to 2015 were used in the study. The data were obtained through BIST web page (<https://datastore.borsaistanbul.com>), from the BIST library and other data sources. The raw data contained 23.222 files belonging to 770 companies listed in BIST.

Since the financial ratio analysis and a bankruptcy prediction in the financial market required different approaches than the non-financial market, it had been a common practice to analyse the bankruptcy in financial and non-financial sectors separately. The analysis in this dissertation study focused on the non-financial companies' bankruptcy. So, the financial sectors (Banks, Insurance Companies, Leasing and Factoring Companies, Holding and Investment Companies, Brokerage Houses, REITs, etc) were needed to be eliminated from the data.

Further, a healthy classification of bankrupt and non-bankrupt companies was required to construct a robust model.

The elimination of companies to obtain a healthy sample was done through several steps.

i. Starting from 769 companies, an initial elimination of financial companies was made.

The market data of the companies in Public Disclosure Platform (PDP) listed after

2009 was present and used for the elimination. 75 companies were eliminated from the list. 694 were selected for further elimination.

- ii. To separate bankrupt and non-bankrupt companies, an assumption was made that if the company reported a financial statement data to BIST in 2015 last quarter, it was a non bankrupt sample candidate, if not, it was a bankrupt candidate. Of 694 companies, 270 didn't report at 2015Q4 so they were candidates for the bankrupt sample.
- iii. Of 694 companies, 474 presented a financial report in 2015 last quarter. So, they were the candidates for the non-bankrupt sample

3.1.1 Bankrupt Companies Selection

270 candidate companies for bankrupt sample were subject to deep, one to one investigation from the BIST website, Public Disclosure Platform, Union of Chambers and Commodity exchanges website, trade registry newspaper site, and several other internet sources, because not all companies, that didn't report financials were bankrupt. In fact, the absence of a report was rarely linked to a bankruptcy. (Table 3.1 below)

- Although a preliminary elimination was made through the market segmentation from PDP data, 113 additional financial services companies were came across and eliminated. They were not eliminated in the prior steps because, the data obtained from PDP that helped the preliminary elimination were available for only the listed companies after 2009. In this step, other financial companies whose markets were not available in PDP website were also eliminated.

- Mergers, Acquisitions and Change of Company Name: 58 companies were merged to or acquired by other companies, or the ticker symbol of the company changed.
- Bankruptcy or Financial Distress: In the notes reported to BIST or PDP, 32 companies discontinued their operations due to financial distress, liquidation, reorganization or bankruptcy. These companies were the main focus area of this study, and their data would be used as the bankrupt data further in these dissertation study.
- Legitimate reasons. 18 companies were confiscated by government agencies (like Savings & Deposits Insurance Fund, etc) after legitimate cases, or cancelling of any licenses.
- Some company filename data were misspelled. 14 companies were eliminated, not to cause a double counting error.
- Continued Operation: Although not presented a report in 2015 Q4, 10 companies sent their reports afterwards, because they were continuing their operations.
- Issue of another Capital Market Instrument: 8 companies sent reports because they were using another market instrument (mostly bonds) and sending yearly reports to the authorities.
- New Companies: 6 companies started operations after 2015 Q4 and started sending reports afterwards.
- Fraud: 5 companies were delisted after news about some fraudulent cases.
- Own will: 5 companies delisted from BIST on their own will. These delisting demands should be in line with the communique of CMB dated 2 Nov 2014 called “Right to squeeze out and sell”

Borsa Istanbul also states the reasons for “Delisting” as

- Negative shareholders’ equity;

- Non-compliance with the rules and regulations, especially violation of disclosure and insider trading regulations;
- Filing for bankruptcy or suspending operations for more than three months;
- Dissolution of the company;
- Company facing difficulties in paying its debts or in financial distress;
- Company not fulfilling the fee obligations of the Exchange;
- Cancellation or invalidation of any permits, licenses or agreements of the company, crucial to its operations;
- An independent auditor gives a negative opinion or refrains from giving an opinion for the last two financial periods;
- Loss of 2/3 of the operational capacity due to any disaster such as fire, war, terrorism, etc

Companies are delisted involuntarily upon the approval of Executive Council through consultation with the Listing Committee.

On the other hand, Borsa Istanbul also lets the companies (excluding investment trusts), whose main shareholders have at least 95% of the capital and/or voting rights, delist voluntarily upon the approval of Capital Markets Board (CMB). The main shareholder(s) should make a tender offer to the minority shareholders. After the tender offer is completed, the company may be delisted.

31 Bankrupt sample companies, their ticker symbols, the period of last reported date and the number of periods reported until bankruptcy are given in Table 3.2 below.

Table 3.1

Reasons for Elimination from the Bankrupt Candidate Sample

Reasons of Elimination	# of companies eliminated
Finance Companies	113
Merger, Acquisition, Change of Name	58
Bankruptcy or Financial Distress	31
Legitimate Reasons	18
Double Counts	16
Continued Operation	10
Issue of another Capital market Instrument	8
New Company	6
Fraud	5
Own Will	5
TOTAL	270

In the model, bankruptcy was to be predicted 2 years before it happened. So, the companies that lived shorter than 2 years and discontinued were out of the focus. For this reason, 4 companies (ARBUL, KOYTS, SOKSA, TEPAS) were removed from the sample of bankrupt companies. The final sample for the bankrupt was consisting of 27 final companies.

In another study about prediction of corporate financial distress by Uğurlu and Aksoy (2006), similar period was addressed for Borsa Istanbul Non-financial Service Companies. The number of bankrupt companies obtained were also very similar. The population of financially distressed companies was 32 and 5 of them were eliminated due to missing data, so the final number of distressed companies were 27 in their study.

3.1.2 Non-bankrupt Companies Selection

424 companies that presented a financial report in 2015Q4 were candidates of non-bankrupt sample. The markets of 132 of them were not available and 43 of them were in financial sectors. So, these 174 companies were eliminated and 249 companies were left for the final non-bankrupt companies' sample (Table 3.3 below).

3.2 Independent variables

Independent variables were selected considering 2 points

- The popularity in the previous bankruptcy literature
- The ability to obtain the related financial accounts from the financial statements of the companies

The unstandardized taxonomy of the financial statements between the companies and between years was a limitation. The same limitation was also mentioned by Uğurlu and Aksoy (2006). To overcome this limitation of unstandardized taxonomy between different reporting years, they focused the analysis on a limited period (1996 – 2003) where the standard was known to be the same.

Table 3.2

Final List of Bankrupt Companies

Company Name	Ticker	Period of Last Report	Number of Reports
Abana Elektromekanik San ve Tic. AŞ	ABANA	2007Q4	40
Apeks Dış Ticaret AŞ	APEKS	2000Q3	11
Arat Tekstil Sanayi ve Ticaret A.Ş.	ARAT	2007Q2	38
Arbul Entegre Tekstil İşletmeleri A.Ş.	ARBUL	2015Q3	7
Ataç İnşaat ve Sanayi A.Ş.	ATAC	2013Q3	10
Berdan Tekstil Sanayi ve Ticaret A.Ş.	BERDN	2013Q1	61
Boyasan Tekstil Sanayi ve Ticaret A.Ş.	BYSAN	2011Q2	48
ÇBS Boya Kimya Sanayi ve Ticaret A.Ş.	CBSBO	2013Q3	63
Favori Dinlenme Yerler A.Ş.	FVORI	2015Q1	60
Goldaş Kuyumculuk Sanayi İthalat İhracat A.Ş.	GOLDS	2012Q1	52
Gorbon Işıl Seramik A.Ş.	GORBN	2003Q1	21
Gümüşsuyu Halı ve Yer Kaplamaları San. ve Tic. A.Ş.	GUMUS	2002Q2	16
Koniteks Konfeksiyon Endüstri Ve Ticaret A.Ş.	KOTKS	2006Q3	35
Köytaş Tekstil San. ve Tic. A.Ş.	KOYTS	2002Q2	2
Lio Yağ Sanayii ve Ticaret A.Ş.	LIOYS	2007Q4	32
Mudurnu Tavukçuluk A.Ş.	MDRNU	2001Q4	16
Meges Boya Sanayi ve Ticaret A.Ş.	MEGES	2008Q3	43
Metaş İzmir Metalurji Fabrikası T.A.Ş.	METAS	2003Q2	21
ÇBS Printaş Oto Boya ve Gereçleri Sanayii A.Ş.	PRTAS	2013Q3	61
Raks Elektronik Sanayi ve Ticaret A.Ş.	RAKSE	2006Q3	30
Ran Lojistik Hizmetleri A.Ş.	RANLO	2013Q1	15
Raks Elektrikli Ev Aletleri San. ve Tic. A.Ş.	RKSEV	2006Q3	35
Sezginler Gıda San. ve Tic. A.Ş.	SEZGD	2002Q3	10
Şeker Pili. Ve Yem Sanayi Ticaret A.Ş.	SKPLC	2014Q3	59
Sinop Örme ve Konfeksiyon Sanayi ve Ticaret A.Ş.	SOKSA	2002Q1	4
Sevgi Özel Sağlık Hizmetleri A.Ş.	SVGSH	2000Q1	9
Tepa Tıbbi ve Elektronik Ürünler Sanayi ve Ticaret A.Ş.	TEPAS	2013Q2	1
Tümteks Tekstil Sanayi ve Ticaret A.Ş.	TUMTK	2011Q2	44
Uki Uluslararası Konfeksiyon İmalat ve Ticaret A.Ş.	UKIM	2008Q4	44
Ünal Tarım Ürünleri İhracat ve Sanayi A.Ş.	UNTAR	2006Q3	34
Uzel Makina Sanayii A.Ş.	UZEL	2009Q4	44

Table 3.3

List of Non-bankrupt Companies

ACSEL	ASUZU	BRKO	DEVA	ESCOM	KAREL	MEPET	PGSUS	SEYKM	TUKAS
ADEL	ATEKS	BRKSN	DGATE	ETILR	KARSN	MERIT	PINSU	SILVR	TUPRS
ADESE	ATPET	BRMEN	DGKLB	FLAP	KARTN	MERKO	PKART	SKTAS	TURGG
AEFES	AVOD	BRSAN	DGZTE	FMIZP	KATMR	METAL	PKENT	SNKRN	ULAS
AFYON	AVTUR	BSOKE	DIRIT	FRIGO	KENT	METUR	PNSUT	SNPAM	ULKER
AKCNS	AYCES	BTCIM	DITAS	FROTO	KERVT	MGROS	POLTK	SODA	ULUSE
AKENR	AYEN	BUCIM	DMSAS	GEDZA	KIPA	MIPAZ	PRKAB	SODSN	ULUUN
AKGUV	AYES	BURCE	DOAS	GENTS	KLMSN	MNDRS	PRKME	TACTR	UNYEC
AKSA	AYGAZ	BURVA	DOBUR	GEREL	KNFRT	MRDIN	PRZMA	TATGD	USAK
AKSEN	BAGFS	CCOLA	DOGUB	GOLTS	KONYA	MRSHL	PSDTC	TBORG	UTPYA
AKSUE	BAKAB	CELHA	DURDO	GOODY	KORDS	NETAS	RODRG	TCELL	UZERB
ALCAR	BALAT	CEMAS	DYOBY	GSDDE	KRONT	NIBAS	ROYAL	TEKTU	VAKKO
ALCTL	BANVT	CEMTS	EDIP	GUBRF	KRSAN	NUHCM	RTALB	TGSAS	VANGD
ALKA	BASCM	CIMSA	EGEEN	HATEK	KRSTL	ODAS	RYSAS	THYAO	VESBE
ALKIM	BEYAZ	CLEBI	EGGUB	HEKTS	KRTEK	OLMIP	SAMAT	TIRE	VESTL
ALYAG	BFREN	CMBTN	EGPRO	HURGZ	KSTUR	ORGE	SANEL	TKNSA	VKING
ANACM	BIMAS	CMENT	EGSER	IHEVA	KUTPO	ORMA	SANFM	TKURU	YAPRK
ANELE	BIZIM	CRFSA	EKIZ	IHGZT	KUYAS	OTKAR	SANKO	TMPOL	YATAS
ANELT	BLCYT	CUSAN	EMKEL	INTEM	LINK	OYLUM	SARKY	TMSN	YAYLA
ARCLK	BMEKS	DAGI	EMNIS	IZFAS	LKMNH	OZBAL	SASA	TOASO	YBTAS
ARENA	BMELK	DARDL	ENKAI	IZMDC	LOGO	OZRDN	SAYAS	TRKCM	YONGA
ARMDA	BNTAS	DENCM	EPLAS	IZOCM	LUKSK	PARSN	SEKUR	TSPOR	YUNSA
ARSAN	BOLUC	DERIM	ERBOS	IZTAR	MAALT	PENGD	SELEC	TTKOM	YYAPI
ASELS	BOSSA	DESA	EREGL	JANTS	MAKTK	PETKM	SELGD	TTRAK	ZOREN
ASLAN	BRISA	DESPC	ERSU	KAPLM	MEGAP	PETUN	SERVE	TUCLK	

Table 3.4

List of Independent Variables

VARIABLE NAME	DEFINITION	VARIABLE TYPE
ACID	$(\text{Current Assets} - \text{Inventories}) / \text{Current Liabilities}$	Scalar
CASHPOS	$(\text{Cash} + \text{Short Term Investments}) / \text{Assets}$	Scalar
CHIN	$(\text{Net income}_t - \text{Net income}_{t-1}) / (\text{Net income}_t + \text{Net income}_{t-1})$	Scalar
CLCA	$\text{Current Liabilities} / \text{Current Assets}$	Scalar
GROSSM	$(\text{Sales} - \text{COGS}) / \text{Sales}$	Scalar
INVTO	$\text{Inventories} / \text{Sales}$	Scalar
LEVER	$\text{Total Debt} / \text{Total Capital}$	Scalar
NITA	$\text{Net Income} / \text{Assets}$	Scalar
SANP	$\text{Sales} / \text{Net Plant}$	Scalar
SIZETRL	$\text{Log (Assets in TRL)}$	Scalar
SIZETRL	$\text{Log (Assets in USD)}$	Scalar
TLTA	$\text{Liabilities} / \text{Assets}$	Scalar
WCTA	$\text{Net Working Capital} / \text{Total Assets}$	Scalar
INTWO_BNR	1 ____ Net Income is Negative for last 2 years 0 ____ Otherwise	Binary /Dummy
OENEG_BNR	1 ____ Liabilities Exceed Assets 0 ____ Otherwise	Binary /Dummy

In this study, to extract the correct accounts from which the financial ratios were calculated, very heavy use of Excel, Visual Basic and other calculations were used. A list of common financial statements' titles in the reports were accumulated separately from different periods that the standards were different. Then a transition file was saved for each company for different financial statements (Balance Sheet and Income Statement). The transition file contained a common financial statement line, the period and the data corresponding that particular item and year for each company. After the construction of these files, master files for each company were saved, with the complete financial data and ratios between 1998 and 2016.

3.3 Selection Bias

There were 2 main sample selection bias existing due to sample selection methodologies which were explained in detail by Zmijewski (1984).

3.3.1 Choice Based Selection Bias

As was explained by Zmijewski, the probability of a firm entering our sample depended on the bankruptcy status of the firm. To mitigate the effects of a choice based sample selection bias, a paired sampling method was not preferred. All of the companies that fulfilled the complete data criterion were chosen to enter into the model.

As we had 276 companies in the sample, of which 27 are elected as bankrupt, the probability of a bankruptcy in the sample was %9,8 ($27 / 276 = \%9,8$). This probability was so much

more than the probability of a bankruptcy mentioned as % 0,75 by Zmijewski. So, although mitigated, there were the effects of a choice based selection in the model.

Zmijewski showed that as the selection of the sample probabilities approached to the probabilities of the population, more correct classification rates were achieved in total classification and in the non-bankrupt classification. On the contrary, as these probabilities converged, less correct classifications of the bankrupt companies got. So, this meant that, as the probability of a bankruptcy was greater in our sample when the population was considered, our model would classify bankrupt companies more accurately. This created an advantage especially when the costs of the errors were considered and when it was known that the error in predicting the bankruptcy was much more costly than the error in predicting the non-bankruptcy.

No specific estimation technique like WESML or any other was used in the model, since the population probabilities were unknown.

3.3.2 Complete Data Criterion

Although a full set of data was not a necessity to enter the sample groups of bankrupt and non-bankrupt companies, the ones with very little reported data were eliminated from the list. For example, as mentioned in Section 3.1.1, four companies of 31 bankrupt were removed from the sample, since they didn't have enough data. So, the probability of a company to enter into our sample was greater, if it was having complete data. Our sample had sample selection bias.

3.4 Outlier Definitions and Data Elimination

Before choosing the proper crosssection for the sample, 2 eliminations were applied

3.4.1 Outliers

The data were grouped according to the sectors of the companies. And the outlier values for each ratio in each sector were selected and eliminated from the data of the non-bankrupt sample as advised by Lev and Sunder (1978). They mentioned two techniques to treat the outliers as trimming and winsorizing. Trimming was the removal of the equal number of outlier points in smallest and largest percentiles. Winsorization was changing the outlier value to a nearest value which was not to be treated as an outlier. The technique applied to eliminate the outliers in this study was trimming the first and last %5 percentiles. After the elimination, the values of each ratio of each sector between the %5 and %95 percentiles were remaining to enter the model.

3.4.2 Economic Crisis Period

The independent variables of the study did not include any macroeconomic indicator. As Stenbäck (2013) mentioned, macroeconomic variables were valuable for prediction models, but their value was observed mostly in the situations of economic shocks. In normal business cycles these variables didn't contribute to default prediction. So, the data for

- 2000 Fourth quarter
- 2001 First quarter
- 2001 Second quarter

- 2001 Third quarter
- 2001 Fourth quarter
- 2002 First quarter
- 2008 Third quarter
- 2008 Fourth quarter
- 2009 First quarter
- 2009 Second quarter
- 2009 Third quarter
- 2009 Fourth quarter

were eliminated from the non-bankrupt sample crossections because, these were the periods of major economic crisis in Turkish Economy. (Türkiye’de 1994, 2001 ve 2008-9 Ekonomik Krizlerinin Karşılaştırmalı Analizi, 2010) (Krizden Krize Türkiye: 2000 Kasım ve 2001 Şubat Krizleri, 2001)

3.5 Data Construction

The model aimed to predict the bankruptcy 2 years before it happened. So, for the bankrupt class sample, the values of each ratio in the financial records corresponding to 2 years before the last record were noted. For different samples constructed, the crossectional data for the bankrupt companies were the same.

The year to pick the crossectional data for the non-bankrupt sample were extracted randomly. A single quarter between 1998 and 2015Q4 was randomly selected for each non-bankrupt company, and the financial ratios belonging to that quarter entered into the sample. So for

different samples constructed, the cross-sectional data for non-bankrupt companies changed in random years.

This method of random selection gave the ability to create multiple samples. In the Findings section, most of the analyses were executed over 5 different samples, in which the dates for the bankrupt companies were belonging to 2 years before the bankruptcy dates and the dates for the non-bankrupt companies were selected randomly as explained here.

3.6 Statistical Method

The statistical method was the logistic regression. Logistic regression is a nonlinear probabilistic regression.

As stated in the Ohlson study, the nonlinear probabilistic model aimed to maximize the $l(\beta)$ (log likelihood function).

$$l(\beta) = \sum_{i \in S_1} \log P(X_i, \beta) + \sum_{i \in S_2} \log (1 - P(X_i, \beta))$$

where

- S_1 is the set of bankrupt firms
- S_2 is the set of non-bankrupt firms
- $l(\beta)$ is the log likelihood of any specific outcomes
- $P(X_i, \beta)$ is the probability of bankruptcy
- X_i is the vector of predictors
- β is vector of parameters

The aim is to find the β parameters where $l(\beta)$ is maximum

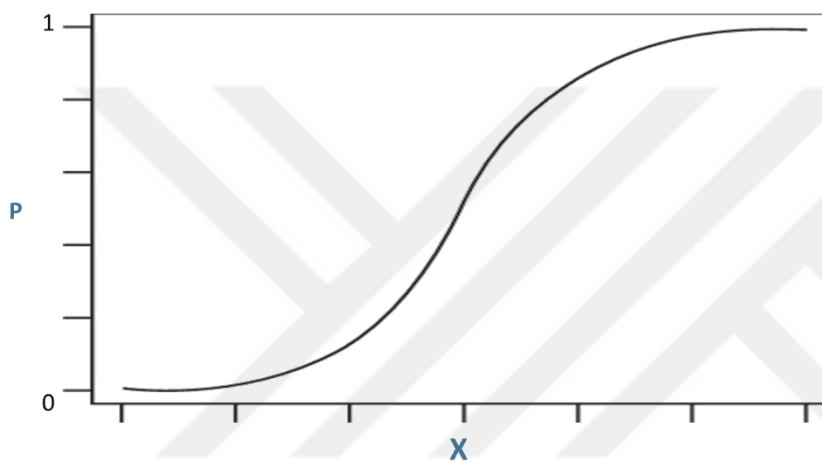
In the logistic regression, P is the sigmoid or logistic function where

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

$$P = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Figure 3.1

Logistic (Sigmoid) Function



4. RESULTS AND FINDINGS

4.1. Mean Values

5 different samples were created in which, the crossection dates for the bankrupt were the same (for each sample, the date 2 years before the bankruptcy were included) and the crossection dates for non-bankrupt were selected randomly, as mentioned above.

The mean values of each of the ratios are displayed in Table 4.1.

The difference of ratios between bankrupt and non-bankrupt companies were significant at %1 level except for ACID, CHIN, GROSSM, SANP and SIZEUSD. (ACID was significantly different at %10 only in one sample, CHIN was different significantly in %1 in 3 samples, and %10 in 2 samples, SIZEUSD was significantly different only in 2 samples in %10). So it could be seen that, the ratios CASHPOS, CHIN, CLCA, INVTO, LEVER, NITA, SIZETRL, TLTA and WCTA differed significantly among bankrupt and non-bankrupt companies.

CASHPOS: Cash Position of the Non-bankrupt companies were significantly greater.

CHIN: CHIN was affected mostly by the decline in the Net Income in the consecutive years. This decline was significantly sharp for the bankrupt companies, when non-bankrupt ones were considered.

CLCA: Bankrupt companies had significantly more current liabilities than their current assets. It was obvious that this pointed out severe destruction in the net working capital.

INVTO: Bankrupt companies were significantly less efficient in converting their inventories to sales.

LEVER: Bankrupt companies used significantly more leverage.

NITA: Bankrupt companies produced negative net profit, significantly lower than the non-bankrupt companies.

SIZETRL: Size effect was observed in the difference of the mean values of the bankrupt and non-bankrupt companies, as frequently observed in the literature. The logarithm of the asset sizes of the non-bankrupt companies were significantly greater than the bankrupt ones.

In this dissertation study, rather than being considered within a paired sample design, Size was included as an independent variable in the model.

It was worth noting that, although SIZETRL variable was significantly different among bankrupt and non-bankrupt, SIZEUSD variable was not significantly different. To calculate the SIZEUSD variable, Asset size of the company in TRL was converted to USD for each asset value. So this meant that, although the asset size in USD was important for companies to show a growth in the asset size, the value in the national currency was more important in determining the size effect.

TLTA: Total of long term and short term liabilities considered with the asset sizes of the bankrupt companies were significantly greater.

WCTA: Net working capital conditions of bankrupt companies were significantly worse in the bankrupt companies.

4.2. Multicollinearity Between Independent Variables

As some of the ratios used common financial accounts in their calculation, they were derivatives of each other, using common components as numerators or denominators. So strong correlation between them were expected. As stated by Courtis (1978), a small group of independent variables were expected to capture most of the information.

The multicollinearities are given from the Sample 1 in the Table 4.2 below. The cross-correlation tables of other samples underlined the same variable partners, having strong correlation between. From Table 4.2, it can be seen that there are high and significant correlations among

- CLCA and NITA
- CLCA and OENEG_BNR
- CLCA and TLTA
- CLCA and WCTA
- OENEG_BNR and TLTA
- OENEG_BNR and WCTA

Table 4.1

Mean Values of the Independent Variables

		Sample1	Sample2	Sample3	Sample4	Sample5
ACID	Nonbankrupt	1.3803	1.3646	1.4779	1.2585	1.2819
	Bankrupt	0.9199	0.9199	0.9199	0.9199	0.9199
	Difference	0.4605	0.4448	0.5580	0.3386	0.3621
CASHPOS	Nonbankrupt	0.0690	0.0668	0.0751	0.0668	0.0737
	Bankrupt	0.0101	0.0101	0.0101	0.0101	0.0101
	Difference	0.0589	0.0567	0.0649	0.0567	0.0636
CHIN	Nonbankrupt	0.0034	-0.0176	0.1062	0.0681	0.0130
	Bankrupt	-0.1838	-0.1838	-0.1838	-0.1838	-0.1838
	Difference	0.1872	0.1661	0.2900	0.2519	0.1968
CLCA	Nonbankrupt	0.6802	0.7432	0.6955	0.7550	0.7134
	Bankrupt	2.8818	2.8818	2.8818	2.8818	2.8818
	Difference	-2.2016	-2.1386	-2.1864	-2.1269	-2.1685
GROSSM	Nonbankrupt	0.2169	0.2078	0.2054	0.2182	0.2210
	Bankrupt	0.1600	0.1600	0.1600	0.1600	0.1600
	Difference	0.0569	0.0478	0.0455	0.0582	0.0610
INVTO	Nonbankrupt	0.3483	0.3576	0.3319	0.3762	0.3682
	Bankrupt	3.3746	3.3746	3.3746	3.3746	3.3746
	Difference	-3.0263	-3.0170	-3.0426	-2.9984	-3.0064
LEVER	Nonbankrupt	1.3311	1.2203	1.2399	1.2860	1.3988
	Bankrupt	4.5117	4.5117	4.5117	4.5117	4.5117
	Difference	-3.1806	-3.2914	-3.2718	-3.2257	-3.1129
NITA	Nonbankrupt	0.0260	0.0260	0.0345	0.0294	0.0275
	Bankrupt	-0.1216	-0.1216	-0.1216	-0.1216	-0.1216
	Difference	0.1476	0.1476	0.1561	0.1510	0.1490
SANP	Nonbankrupt	9.8579	9.6097	9.1871	7.7794	4.7969
	Bankrupt	9.9396	9.9396	9.9396	9.9396	9.9396
	Difference	-0.0817	-0.3299	-0.7525	-2.1602	-5.1427
SIZETRL	Nonbankrupt	18.3639	18.4345	18.4410	18.4684	18.4034
	Bankrupt	17.5653	17.5653	17.5653	17.5653	17.5653
	Difference	0.7985	0.8692	0.8757	0.9030	0.8380
SIZEUSD	Nonbankrupt	17.8415	17.8210	17.8770	17.8824	17.8495
	Bankrupt	17.3994	17.3994	17.3994	17.3994	17.3994
	Difference	0.4420	0.4216	0.4775	0.4830	0.4500
TLTA	Nonbankrupt	0.4630	0.4674	0.4796	0.4819	0.4804
	Bankrupt	1.3224	1.3224	1.3224	1.3224	1.3224
	Difference	-0.8594	-0.8551	-0.8429	-0.8405	-0.8420
WCTA	Nonbankrupt	0.1781	0.1631	0.1826	0.1650	0.1745
	Bankrupt	-0.4448	-0.4448	-0.4448	-0.4448	-0.4448
	Difference	0.6229	0.6079	0.6274	0.6099	0.6194

Significant at %1 level***Significant at %10 level***

Table 4.2

Multicollinearity Between Independent Variables

Correlation Probability	ACID	CASHPOS	CHIN	CLCA	GROSSM	INVTO	LEVER	NITA	OENEG_BNR	SANP	SIZETRL	SIZEUSD	TLTA	WCTA
ACID	1													

CASHPOS	0.215	1												
	0.001	-----												
CHIN	0.135	0.084	1											
	0.083	0.275	-----											
CLCA	-0.253	-0.184	-0.142	1										
	0.000	0.005	0.069	-----										
GROSSM	0.214	0.145	0.228	-0.049	1									
	0.002	0.031	0.003	0.476	-----									
INVTO	-0.060	-0.087	0.064	0.407	0.264	1								
	0.377	0.198	0.419	0.000	0.000	-----								
LEVER	-0.147	-0.108	-0.004	-0.052	0.013	-0.075	1							
	0.025	0.098	0.955	0.431	0.848	0.270	-----							
NITA	0.252	0.308	0.236	-0.705	0.045	-0.392	-0.076	1						
	0.000	0.000	0.002	0.000	0.523	0.000	0.275	-----						
OENEG_BNR	-0.201	-0.318	0.186	0.736	0.317	0.502	-0.412	-0.516	1					
	0.305	0.099	0.344	0.000	0.107	0.008	0.029	0.005	-----					
SANP	-0.017	-0.061	0.058	-0.023	-0.067	-0.021	0.021	0.015	-0.115	1				
	0.806	0.361	0.465	0.741	0.329	0.761	0.760	0.834	0.561	-----				
SIZETRL	-0.041	0.035	0.073	-0.130	-0.021	-0.131	-0.028	0.207	-0.184	-0.011	1			
	0.539	0.598	0.344	0.051	0.754	0.056	0.676	0.003	0.348	0.878	-----			
SIZEUSD	-0.035	0.031	0.083	-0.123	0.011	-0.121	-0.005	0.186	-0.292	-0.009	0.967	1		
	0.595	0.637	0.283	0.066	0.866	0.077	0.945	0.007	0.132	0.898	0.000	-----		
TLTA	-0.316	-0.186	0.010	0.764	0.033	0.400	0.104	-0.574	0.866	0.016	-0.118	-0.107	1	
	0.000	0.004	0.894	0.000	0.625	0.000	0.103	0.000	0.000	0.817	0.073	0.105	-----	
WCTA	0.269	0.200	0.109	-0.861	0.010	-0.322	0.019	0.640	-0.754	0.044	0.148	0.135	-0.882	1
	0.000	0.002	0.160	0.000	0.887	0.000	0.772	0.000	0.000	0.514	0.027	0.043	0.000	-----

- SIZETRL and SIZEUSD
- TLTA and WCTA

These multicollinear values were not surprising, since they were mostly derivatives of each other, as stated above.

So, the independent variables

- CLCA
- OENEG_BNR
- SIZEUSD and
- TLTA were to be removed from the logistic model.

4.3. Construction of the Logistic Model

4.3.1. Stepwise Regression

The independent variables that didn't have multicollinearity among were introduced into the stepwise logistic regression. These variables were ACID, CASHPOS, CHIN, GROSSM, INVTO, LEVER, NITA, SANP, SIZETRL, WCTA.

In the Table 4.3 below, the coefficients and p values of the independent variables in the first step of the logistic regression are given.

For CASHPOS, CHIN, NITA, SIZETRL and WCTA, the coefficients were negative and for INVTO and LEVER, the coefficients were positive in all of the samples. Whereas for ACID,

GROSSM and SANP the direction of the coefficients was not definite. This points out the same inference given in Table 4.1 in which, the difference of ratios of bankrupt and non-bankrupt companies were given. (It can be remembered that in Table 4.1, the differences in ACID, GROSSM and SANP were not significantly different between bankrupt and non-bankrupt)

Table 4.3

Signs of the Independent Variables

	Sample1		Sample2		Sample3		Sample4		Sample5		Signs
	Coefficient	p value	Coefficient	p value	Coefficient	p value	Coefficient	p value	Coefficient	p value	
C	19.9720	0.0323	3.6409	0.4740	7.3446	0.2745	8.6499	0.1826	10.5984	0.1178	+
ACID	-0.2976	0.6407	0.2031	0.6384	0.0849	0.8907	0.0643	0.8892	0.4402	0.5793	nd
CASHPOS	-34.1596	0.0657	-45.7315	0.0511	-47.8279	0.0302	-82.1863	0.0112	-47.3187	0.1518	-
CHIN	-0.4136	0.8264	-1.8156	0.1171	-1.3503	0.4367	-2.7457	0.0229	-2.1735	0.1833	-
GROSSM	-18.8937	0.0349	1.1603	0.7489	2.3604	0.2483	0.0350	0.9950	-2.3024	0.6948	nd
INVTO	2.7468	0.1720	0.1287	0.6903	0.7439	0.6835	0.1094	0.6455	0.0898	0.7603	+
LEVER	0.2880	0.3003	0.3246	0.0415	0.4617	0.1163	0.4548	0.0545	0.5781	0.1247	+
NITA	-27.1647	0.0869	-9.7131	0.1726	-19.4233	0.1384	-13.8684	0.2367	-12.3700	0.1290	-
SANP	-0.0026	0.5039	0.0233	0.2181	0.1567	0.0613	0.0228	0.0830	0.0727	0.1616	nd
SIZETRL	-0.9443	0.0302	-0.2893	0.2622	-0.5090	0.1606	-0.5008	0.1191	-0.6323	0.0598	-
WCTA	-4.5175	0.2526	-1.4613	0.4282	-6.9120	0.1322	-2.7285	0.4862	-10.1595	0.0549	-

In the stepwise logistic regression, the independent variable with the highest p value (lowest significance) was eliminated and the stepwise process was repeated with the remaining independent variables, until each of the variables had p values significant at %5 level.

Table 4.4

Coefficients of the Final Model for Each Sample

	Sample1		Sample2		Sample3		Sample4		Sample5	
	Coefficient	p value	Coefficient	p value	Coefficient	p value	Coefficient	p value	Coefficient	p value
C	-1.594077	0.0015	-1.610243	0.0008	-1.37465	0.0063	-1.230728	0.0248	-1.800798	0.0001
CASHPOS	-41.23956	0.0135	-50.96819	0.0153	-55.18121	0.0134	-67.82803	0.0032	-34.98102	0.0197
LEVER	0.412776	0.0134	0.371506	0.009	0.414251	0.0136	0.472047	0.0123	0.357528	0.0134
NITA	-19.53341	0.0229	-23.0275	0.0002	-26.40599	0.0001	-26.55119	0.0002	-24.93966	0.0003
WCTA	-4.114127	0.0619								

In the final model of each sample, CASHPOS, LEVER and NITA were significant in all, WCTA was significant in only Sample 1. (Table 4.4)

So, it was proven that,

- Cash Position of a company was an important determinator of a bankruptcy. The sign of the coefficient showed that, the stronger Cash Position of the company, the less probable a bankruptcy was.
- As proven by many of the previous literature, Leverage was another factor, influencing the bankruptcy prospect of the company. The positive sign of the coefficient showed that, more levered the company was, more probable a bankruptcy was.
- Profitability of the company was another strong determinator of the bankruptcy. The negative coefficient showed that, as the profitability of the company increased, the probability of a bankruptcy was decreasing.
- The significant presence of WCTA in one of the samples showed that net working capital had effect in determining a bankruptcy. The negative sign showed that, if the company had more net working capital considered with its assets, a bankruptcy was less probable.
- The negative sign of the constant favoured that, given no data (about any financial status of the company), it was more probable that it would be classified as non-bankrupt. That was because, we were able to say that, without any financial data and any statistical model, we had 27 bankrupt and 249 non-bankrupt companies. So in a random guess, the bankruptcy probability was $27 / 276$, whereas the non-bankruptcy probability was $249 / 276$.

In a bankruptcy prediction study by Vuran (2009), the Turkish manufacturing companies were studied for the period of 1997 – 2007. Similar variables were found to be significant after their logistic regression. It was found that

- the total asset profitability,
- the financial leverage and
- the measures of liquidity (Cash Flow from Operations / Interest Expense)

were significant in predicting bankruptcy two years prior.

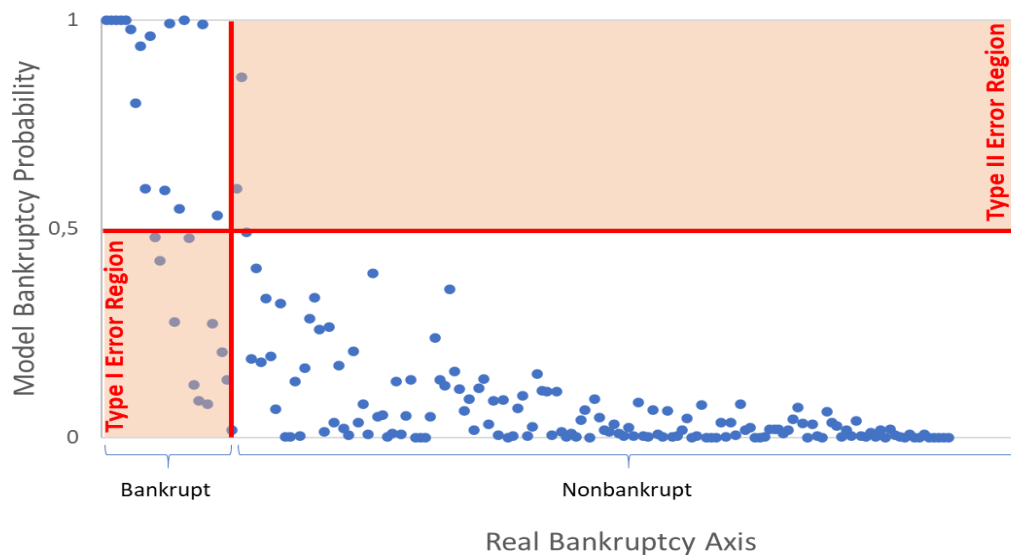
4.3.2. Classification Accuracies

The output of the model was evaluated according to a certain threshold probability called the cut-off point. If the output probability was greater than the cut-off point, the company was classified as bankrupt. If a probability lower than the cut-off was obtained, it was classified as non-bankrupt.

When each company was displayed in a scatter plot with its real and predicted bankruptcy attributes, Type I and Type II Regions were easily visualised

From the plot in Figure 4.1 below, it is also clear that, the number of companies in 4 regions are depending onto the cut-off point. To illustrate, if the cut-off point is chosen as %0, than this means that, all of the companies would be classified as bankrupt. This would result a %100 accurate prediction of bankrupt companies, but no insight about the non-bankrupt group. On the other edge, if the cut-off is chosen as %100, than all of the companies are expected to be categorized as non-bankrupt. As shown in the Figure 4.1 and taken reference in Table 4.5, the common practice in classification was to choose the cut-off point as %50. So, the bankrupt companies with model output less than %50 probability were displayed in the Type I error region, likewise, the non-bankrupt companies with model output greater than %50 probability s displayed in the Type II error region.

Figure 4.1

Type I and Type II Error Regions Representation

For each sample, Expectation – Prediction Evaluation Tables were produced in E-views.

These tables gave the cross tabulations of the bankrupt and non-bankrupt companies in real and model dimensions. Tables were given for each sample in Table 4.5. E-views evaluates the companies with complete independent variables. So although 276 companies entered into each sample (27 bankrupt and 249 non-bankrupt), less amount of them were presented in these cross tabs, since the companies with empty or null values were not entering into the evaluation in these tables.

The goodness of fit was measured with 2 methods.

- i. The gain of the model compared with a random guess are displayed in Table 4.5. For example, without a model, the probability of choosing a bankrupt company was $27 / 183 = \%14,7$. The output of the model pointed 20 companies for bankrupt class, of which, 18 was bankrupt in the real sample also. So the prediction accuracy of the

model in the bankrupt class was %90 (18 / 20) . That meant that, the model produced an accuracy in the bankrupt class, which produced 6,1 times more accurate result than a random guess.

Although it was clear that the model offered improved precision in bankruptcy prediction, there were still certain bankrupt companies that were classified as non-bankrupt by the model. Type I Error Rate* was the conditional probability of a company to be classified as non-bankrupt, given it was bankrupt in real. (This was the definition of PI, conditional probability of the Type I error in Koh Model). In the same context, Type II Error Rate* was the conditional probability of a company to be classified as bankrupt, given it was not, in real (PII in Koh model).

Type I Error Rate and Type II Error rate were the ratios of the number of companies in the Type I and Type II error regions to the total number of companies, respectively.

Total Prediction Accuracy was the number of correctly classified companies over the total number of companies. So, the Total Prediction Accuracy + Type I Error Rate + Type II Error rate equal to 1.

ii. The ROC (Receiver Operator Characteristic) Curve and AUC (Area Under the Curve)

The ROC Curve gives us the ability to compare the goodness of fit of several different classification models.

Table 4.5

Classification Accuracies

SAMPLE I					SAMPLE II				
		MODEL CLASS		TOTAL			MODEL CLASS		TOTAL
		Bankrupt	Nonbankrupt				Bankrupt	Nonbankrupt	
REAL	Bankrupt	18	9	27	REAL	Bankrupt	17	10	27
CLASS	Nonbankrupt	2	154	156	CLASS	Nonbankrupt	5	159	164
	TOTAL	20	163	183		TOTAL	22	169	191
Total Prediction Accuracy					94.0%				
Bankrupt Prediction Accuracy					66.7%				
NonBankrupt Prediction Accuracy					98.7%				
Gain (For Bankruptcy Prediction)					6.10				
Type I Error Rate*					33.3%				
Type I Error Rate					4.9%				
Type II Error Rate*					1.3%				
Type II Error Rate					1.1%				
SAMPLE III					SAMPLE IV				
		MODEL CLASS		TOTAL			MODEL CLASS		TOTAL
		Bankrupt	Nonbankrupt				Bankrupt	Nonbankrupt	
REAL	Bankrupt	18	9	27	REAL	Bankrupt	18	9	27
CLASS	Nonbankrupt	4	160	164	CLASS	Nonbankrupt	2	156	158
	TOTAL	22	169	191		TOTAL	20	165	185
Total Prediction Accuracy					93.2%				
Bankrupt Prediction Accuracy					66.7%				
NonBankrupt Prediction Accuracy					97.6%				
Gain (For Bankruptcy Prediction)					5.79				
Type I Error Rate*					33.3%				
Type I Error Rate					4.7%				
Type II Error Rate*					2.4%				
Type II Error Rate					2.1%				
SAMPLE V									
		MODEL CLASS		TOTAL			MODEL CLASS		TOTAL
		Bankrupt	Nonbankrupt				Bankrupt	Nonbankrupt	
REAL	Bankrupt	16	11	27					
CLASS	Nonbankrupt	2	146	148					
	TOTAL	18	157	175					
Total Prediction Accuracy					92.6%				
Bankrupt Prediction Accuracy					59.3%				
NonBankrupt Prediction Accuracy					98.6%				
Gain (For Bankruptcy Prediction)					5.76				
Type I Error Rate*					40.7%				
Type I Error Rate					6.3%				
Type II Error Rate*					1.4%				
Type II Error Rate					1.1%				

Type I Error Rate* is calculated as in the Koh Model. Number of companies in the Type I Error region/ Number of Bankruptcies

Type II Error Rate* is calculated as in the Koh Model. Number of companies in the Type II Error region/ Number of Nonbankruptcies

Type I Error Rate is calculated as Number of companies in the Type I Error region / Number of All Companies

Type II Error Rate is calculated as Number of companies in the Type II Error region / Number of All Companies

		Model Class	
		Bankrupt	Non-bankrupt
Actual Class	Bankrupt	True Positives	False Negatives
	Non-bankrupt	False Positives	True Negatives

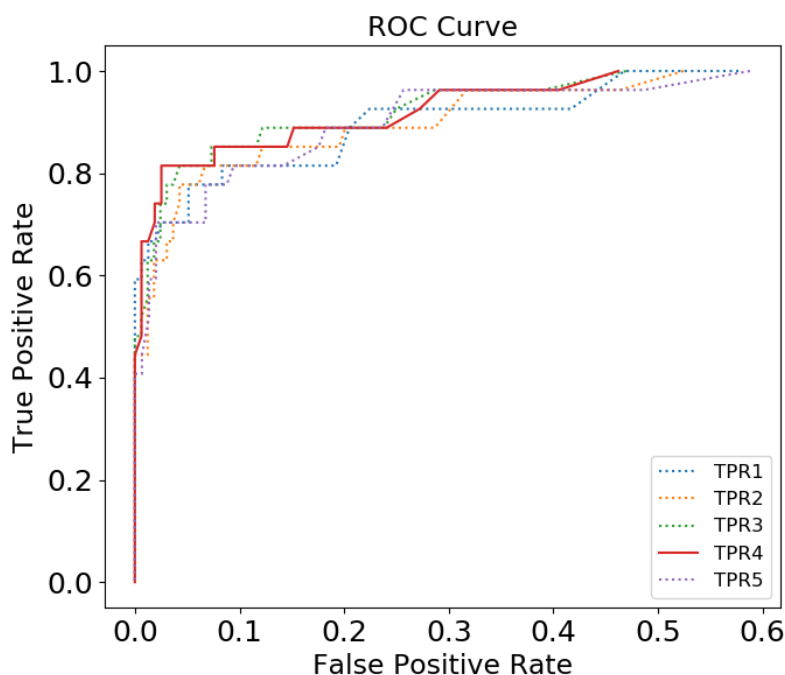
$$\text{True Positive Rate} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{False Positive Rate} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}}$$

The ROC Curves were drawn and compared by changing the cut-off point in each sample model from %0 to %100. From the ROC Curves (Figure 4.2), it was not definite which model fitted its sample best. So, the AUC value (Areas Under the ROC Curve) was calculated to give the answer.

Figure 4.2

The ROC Curves



The AUC values for each model were as

Table 4.6

AUC Values of Different Samples

	AUC Value
Sample1	63.78
Sample2	60.62
Sample3	65.26
Sample4	66.89
Sample5	60.52

So, from the gain displayed in the previous section, and AUC values presented, Model 4 was chosen to have the best goodness of fit. (Model 4 had also been tested to give the greatest AUC value on some other randomly generated samples)

$$P = \frac{1}{1 + e^{-\{-1.230728 + (-67.82803 \times \text{CASHPOS}) + (0.472047 \times \text{LEVER}) + (-26.55119 \times \text{NITA})\}}}$$

4.4. Application of the Koh's Expected Misclassification Cost Model

Koh introduced the Expected Misclassification Cost Model as

$$EC = P_B \times P_I \times C_I + P_N \times P_{II} \times C_{II}$$

where

- EC : Expected misclassification cost of using the model
- P_B : Prior probability of bankruptcy
- P_N : Prior probability of non-bankruptcy
- P_I : Conditional probability of Type I errors = (number of Type I errors given a particular cut-off point)/(number of bankruptcies)

- P_{II} : Conditional probability of Type II errors = (number of Type II errors given a particular cut-off point)/(number of bankruptcies)
- C_I : Misclassification cost of a Type I error
- C_{II} : Misclassification cost of a Type II error

The prior probabilities of bankruptcy (P_B) and non-bankruptcy (P_N) were known a-priori. In the Private Sector Data on Business Failures (Duncan, 1983), the failure rates of sectors excluding finance, insurance, real estate, amusement and recreation services, health services, legal services, educational services, social services, museums, membership organizations, private households, and miscellaneous services were given historically. Most recent records of failures were given as 42 and 62 failures per 10.000 companies for 1980 and 1981 respectively. So P_B and P_N may be considered as %0.5 and %99.5 respectively.

The misclassification costs C_I and C_{II} were deeply analysed in Altman ZETA study. In the related study, C_I / C_{II} was calculated as 35 times.

The expected misclassification cost was a function of the cut off point defined, since the conditional probabilities of P_I and P_{II} were calculated according to these cut-off points.

$$EC = P_B \times P_I \times C_I + P_N \times P_{II} \times C_{II}$$

$$EC = \%0.5 \times P_I \times 35C_{II} + \%99.5 \times P_{II} \times C_{II}$$

So the estimated misclassification cost of misclassification may be standardized by the value of C_{II} as

$$EC^* = \frac{EC}{C_{II}}$$

$$= \%0.5 \times P_I \times 35 + \%99.5 \times P_{II}$$

EC* for each sample simulated at cut-off point 0.5 is given in Table 4.7 as

Table 4.7

Estimated Misclassification Costs of Different Samples

	EC*
Sample 1	0.07108974
Sample 2	0.09515018
Sample 3	0.08260163
Sample 4	0.07092827
Sample 5	0.08474224

So, it could be stated that, in a total Turkish nonfinancial sectors perspective, the expected cost of misclassification in the bankruptcy prediction was

$$EC = EC^* \times C_{II}$$

EC = %7 - %10 times the misclassification cost in the Type II Error Region.

Since C_{II} was mostly defined as an opportunity cost, the lost opportunity of an investor on an investment and get the dividend income or a capital gain, and most probably substituted with another investment of the risk level, the expected misclassification cost in this total perspective was trivial.

Since the aim of this study was not the exact calculation of this misclassification cost, the calculation of the opportunity lost and the misclassification cost mentioned was not detailed further. This subject may be a discussion point for further study.

But as the aim of the study was analysing the change of the misclassification cost around the cut-off point, the question arises: Is the cut-off point of %50 optimal when different costs in the error regions were considered?

To answer this question, EC* was calculated for different combinations of C_I / C_{II} (instead of 35 times) to represent relatively different costs in Type I and Type II error regions. C_I / C_{II} is represented as N

$$EC = P_B \times P_I \times C_I + P_N \times P_{II} \times C_{II}$$

$$N = \frac{C_I}{C_{II}}$$

$$EC = P_B \times P_I \times N \times C_{II} + P_N \times P_{II} \times C_{II}$$

So we standardize EC again as

$$EC^* = \frac{EC}{C_{II}}$$

$$EC^* = P_B \times P_I \times N + P_N \times P_{II}$$

And as was referred to the study by (Duncan, 1983) above, P_B and P_N were calculated as %0.5 and %99.5 respectively again.

$$EC^* = \%0.5 \times P_I \times N + \%99.5 \times P_{II}$$

EC* was tabulated in Figure 4.3 for different N and different cut-off point (meaning different P_I and P_{II})

The results are displayed for Sample 4. The same results are observed for the other 4 samples.

From the representation, it can clearly be seen that, for different N , the optimal cut-off point minimizing the Expected Misclassification Cost was changing.

To look another way, in a 3D perspective, the relation between EC^* , N and cut-off point is displayed. From the 3D graph (Figure 4.4), it can be concluded that, if the optimal cut-off point was not assumed to change for different cut-off points, then the 3D graph should have a planar view. But it can be seen that, the 3D view represented a surface that has different estimated misclassification cost characteristics for different N and different cut-off points.

Figure 4.3

Expected Misclassification Costs vs Different Cut-offs

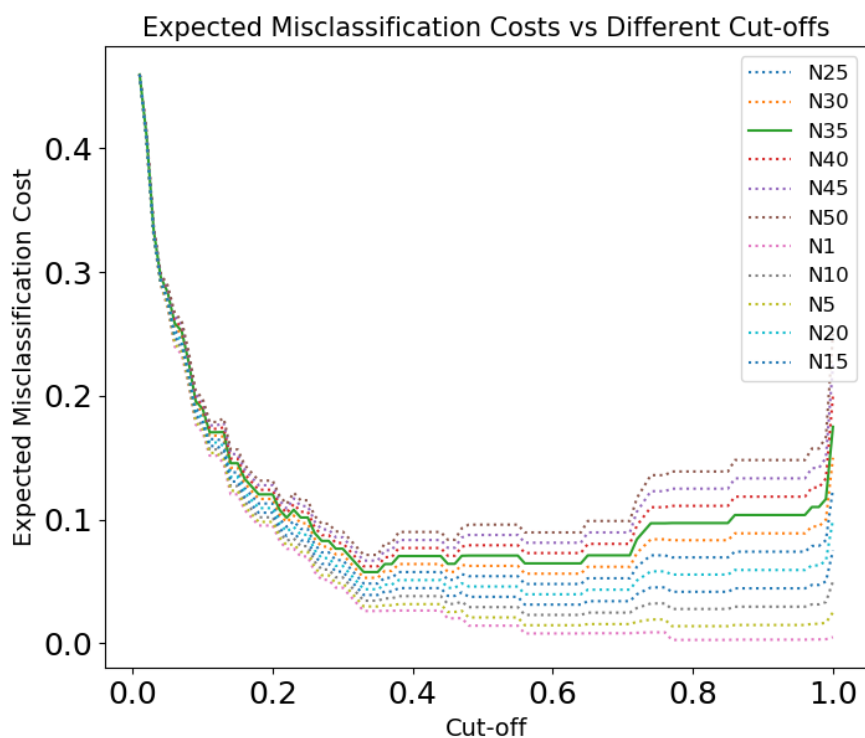
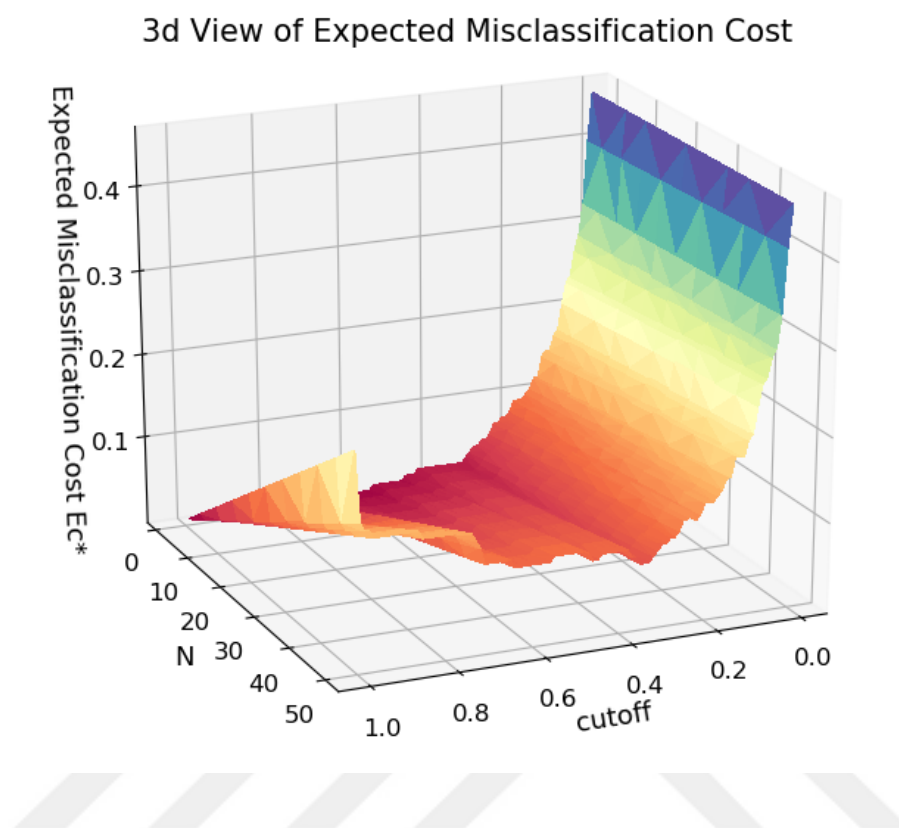


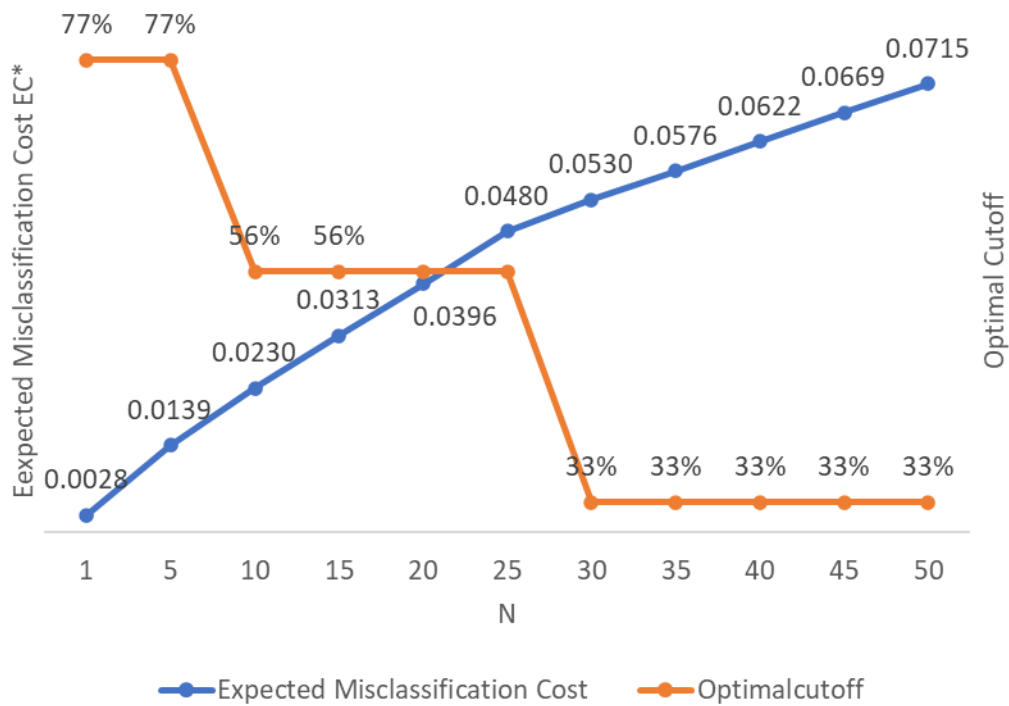
Figure 4.4

3D View of Expected Misclassification Costs

What were the values of minimized expected misclassification costs at different N's and what were the values of the cut-off points to give these minimized cost levels?

Sample 4 is illustrated again in Figure 4.5. Until $N = 10$, the cut-off point that minimized the expected misclassification cost was %77. Between $N=15$ and $N = 25$, a sharp decline to a cut-off point of %56 was required and after $N=25$, a cut-off of %33 was optimal. The minimized expected (standardized) misclassification cost was also displayed here. In Table 4.7, the misclassification cost for $N= 35$ at cut-off point %50 was given as 0.0709. Here it can be seen that, the expected misclassification cost was more optimal (0.0576), at a different cut-off point of %33.

Figure 4.5

Optimal Cut-offs and Minimized Expected Misclassification Costs

When the cut-off point to minimize the expected misclassification and maximize the number of companies correctly classified were compared, they were not equal for each N. Since it was clear that, the cut-off point to maximize the number of correctly classified companies were independent of the costs related to misclassification, it was constant for every N, whereas, as shown in Figure 4.5 and Table 4.8, the cut-off point to minimize the misclassification cost was different for different N.

On the other hand, for $N > 30$, these two cut-off points were not different.

Table 4.8

Expected Misclassification Costs vs Different Cut-offs

N	Cut off to minimize the expected misclassification cost	Cut off to maximize the number of correctly classified companies
1	%77	
5	%77	
10	%56	
15	%56	
20	%56	
25	%56	%33
30	%33	
35	%33	
40	%33	
45	%33	
50	%33	

5. LIMITATIONS AND FURTHER RESEARCH

As mentioned by Kahya, the legal definition of failure might be somehow contaminated, resulting in false indication of healthy and failed companies. In Turkey, some of the companies that had good operational results stopped its operations and declared bankruptcy because of legitimate reasons. On the contrary, some companies that had poor operational results were survivors with unknown ability. This was a limitation for the choice based sampling of the study

The sample was chosen from the Borsa İstanbul listed companies, with an expectation of a standard to reach accurate data about dates of bankruptcies, sectors and financial statement data. But this was not the case. There was a lack of standardization between different time periods and different companies. This was another important limitation for the study.

There might be different inflationary effects on the accounts building up the ratios. These inflationary effects were neglected because of the usage of the ratios instead of the accounts themselves. In other words, the inflationary effects were assumed to be the same for all of the accounts that built up the ratios and to be mitigated by the usage of the ratio form. This might be a limitation for this study and a further research taking these inflationary effects into consideration might be conducted.

Further research may be conducted by using several contents of this study. Rather than the prediction of bankruptcy, the model constructed in the study can be used to rank the companies, and this kind of a ranking can be used for buying or selling decisions, and the

performance of such a portfolio shaped with these buying and selling decisions can be compared with the performances of several market portfolios.

The significant independent variables of the model used in a decision-making algorithm like TOPSIS can be another topic for further study.

The improvement of the model in Turkish Market by the usage of advanced machine learning and artificial intelligence tools can also be an area for research for data scientists.



6. CONCLUSION

This study aimed to find, if the cut-off point defined in the probabilistic models was to be changed, when the costs related to bankruptcy were considered.

Logistic model construct was one of the outputs of the study. Although the model could be subject to change for different time periods, the independent variables of the model that significantly affected bankruptcy were the variables of liquidity, leverage position and profitability. So, this study confirmed and acknowledged the frequent use of these ratios by business practitioners and academicians to find the financial position of a company.

Most of the probability models used %50 as a general cut-off point. If the output of the probability model was greater than %50, than the company was classed as bankrupt.

In a general perspective, the probability of a bankruptcy was very rare. According to Private Sector Data on Business Failures by Joseph W. Duncan of Dun & Bradstreet, the probability of a bankruptcy was %0.5 . On the other hand, the costs of making errors in different error regions were also very different. CI / CII (Costs related to Type I and Type II Errors) was estimated to be 35 times. This meant that, although bankruptcy was very rare, when it came into question, the costs related to it were nontrivial.

From a geometric viewpoint, if the expected misclassification costs were not changing, then the 3D representation of these costs with respect to different cut-off points and different CI /

CII ratios was expected to have a planar view. But it hadn't. The cut-off point to minimize the expected cost was not the same for different cost characteristics and it was different than %50.

The cut-off point to maximize the number of correctly classified companies was not always equal to the cut-off point to minimize the expected misclassification costs. But for the CI / CII ratios greater than 30, they were the same. So, it also could be concluded that for the CI / CII levels mentioned in literature, these two cut-off points were the same, but not equal to %50.



7. REFERENCES

- Aharony, J., Jones, C. P., & Swary, I. (1980). An analysis of risk and return characteristics of corporate bankruptcy using capital market data. *The Journal of Finance*, 35(4), 1001 - 1016.
- Altaş, D., & Giray, S. (2005). Mali Başarısızlığın Çok Değişkenli İstatistiksel Yöntemlerle Belirlenmesi: Tekstil Sektörü Örneği. *Sosyal Bilimler Dergisi*(2), 13 - 28.
- Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(No:4), pp:589 - 609.
- Altman, E. (1980, November). Commercial Bank Lending: Process, Credit Scoring, and Costs of Errors in Lending. *Journal of Financial and Quantitative Analysis*, 15(4), 813 - 832.
- Altman, E. (1984). A Further Empirical Investigation of the Bankruptcy Cost Question. *The Journal of Finance*, 39(4), 1067 - 1089.
- Altman, E., Haldeman, R., & Narayanan, P. (1977). ZETA analysis, a new model for bankruptcy classification. *Journal of Banking and Finance*, 29 - 54.
- Aziz, A., & Lawson, G. H. (1989). Cash Flow Reporting and Financial Distress Models: Testing of Hypotheses. *Financial Management*, 18(1), 55-63.
- Aziz, A., Emanuel, D. C., & Lawson, G. H. (1988 September). Bankruptcy Prediction-An Investigation of Cash Flow Based Models. *Journal of Management Studies*, 25(5), 419-437.
- Barnes, P. (1987). The Analysis and Use of Financial Ratios: A Review Article. *Journal of Business Finance and Accounting*, Winter 14(4), 449 - 461.
- Beaver, W. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, 4, pp. 71-111.
- Beaver, W. (1968). Market Prices, Financial Ratios, and the Prediction of Failure. *Journal of Accounting Research*, 6(2), 179-192.
- Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies*, 267 - 284.
- Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A Review of Bankruptcy Prediction Studies: 1930 to Present. *Journal of Financial Education*, 33, 1-42.
- Blum, M. (1974, Spring). Failing Company Discriminant Analysis. *Journal of Accounting Research*, 12(1), 1 - 25.

- Casey, C., & Bartczak, N. (1985, Spring). Using Operating Cash Flow Data to Predict Financial Distress: Some Extensions. *Journal of Accounting Research*, 23(1), 384-401.
- Casey, C., McGee, V., & Stickney, C. (1986, April). Discriminating between Reorganized and Liquidated Firms in Bankruptcy. *The Accounting Review*, 61(2), 249-262.
- Chen, H.-J., Huang, S. Y., & Lin, C.-S. (2009). Alternative diagnosis of corporate bankruptcy: A neuro fuzzy approach. *Expert Systems with Applications*, 36(4), 7710-7720.
- Chen, M.-Y. (2011). Predicting corporate financial distress based on integration of decision tree classification and logistic regression. *Expert Systems with Applications*, 38(9), 11261-11272.
- Clark, T. A., & Weinstein, M. I. (1983). The Behavior of the Common Stock of Bankrupt Firms. *The Journal of Finance*, 38(2), 489 - 504.
- Coşkun, E., & Sayılğan, G. (2008). Indirect Costs of Financial Distress: Evidence from ISE Listed Companies. *Gazi Üniversitesi, İktisadi ve İdari Bilimler Dergisi*, 10(3), 45 - 66.
- Courtis, J. (1978). Modelling a Financial ratios Categorical Framework. *Journal of Business Finance and Accounting*, 371 - 386.
- Dambolena, I. G., & Shulman, J. M. (1988). A Primary Rule for Detecting Bankruptcy: Watch the Cash. *Financial Analysts Journal*, 44(5), 74 - 78.
- Deakin, E. B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 10(1), 167 - 179.
- Dopuch, N., Holthausen, R. W., & Leftwich, R. W. (1987). Predicting Audit Qualifications with Financial and Market Variables. *The Accounting Review*, 62(3), 431 -454.
- Duncan, J. W. (1983). Private Sector Data on Business Failures. *Review of Public Data Use*, 29-35.
- Eisenbeis, R. (1977). Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics. *The Journal of Finance*, 32(3), 875 - 900.
- Ekşi, İ., & Akçi, Y. (2009). The Effects of Sectoral Differences on Financial Rates: An Empirical Study on IMKB Manufacturing Industry. *The Journal of Faculty of Economics and Administrative Sciences*, 14, Pp:115 - 126.
- Fitzpatrick, P. J. (1934). Transitional Stages of a Business Failure. *The Accounting Review*, 9(4), 337 - 340.
- Gentry, J., Newbold, P., & Whitford, D. (1985). Classifying Bankrupt Firms with Funds Flow Components. *Journal of Accounting Research*, 23(1), 146-160.
- Gombola, M. J., Haskins, M. E., Ketzi, E., & Williams, D. D. (1987). Cash Flow in Bankruptcy Prediction. *Financial Management*, 16(4), 55-65.
- Gupta, M., & Huefner, R. (1972). A Cluster Analysis Study of Financial Ratios and Industry Characteristics. *Journal of Accounting Research*, 10(1), 77-95.

- Gurel, O. (2010). Türkiye’de 1994, 2001 ve 2008-9 Ekonomik Krizlerinin Karşılaştırmalı Analizi. *Ekonomik Yaklaşım*, 21(75), 27-75.
- Gümüş, T., Köroğlu, Ç., & Bulak, Y. (2018). Investigation of Financial Performance of Businesses Operating in BIST30 Industry Index. *International Journal of Disciplines Economics&Administrative Sciences Studies*, 4(8), Pp:286 - 304.
- Haugen, R. A., & Senbet, L. W. (1978). The Insignificance of Bankruptcy Costs to the Theory of Optimal Capital Structure. *The Journal of Finance*, 33(2), 383 - 393.
- Hopwood, W., McKeown, J., & Mutchler, J. (1989). A Test of the Incremental Explanatory Power of Opinions Qualified for Consistency and Uncertainty. *The Accounting Review*, 64(1), 28-48.
- Horrigan, J. (1965, July). Some Empirical Bases of Financial Ratio Analysis. *The Accounting Review*, 40(3), 558 - 568.
- İçerli, Y., & Akkaya, C. (2006). Finansal Açıdan Başarılı Olan İşletmelerle Başarısız Olan İşletmeler Arasında Finansal Oranlar Yardımıyla Farklılıkların Tespiti. *İktisadi ve İdari Bilimler Dergisi*, 20, Pp:413 - 421.
- Joy, M., & Tollefson, J. (1975). On the Financial Applications of Discriminant Analysis. *The Journal of Financial and Quantitative Analysis*, 10(5), 723 - 739.
- Kahya, E., & Theodossiou, P. (1999). Predicting Corporate Financial Distress: A Time-Series CUSUM Methodology. *Review of Quantitative Finance and Accounting*(13), 323-345.
- Kalaba, R. E., Langetieg, T. C., Rasakhoo, N., & Weinstein, M. I. (1984). Estimation of Implicit Bankruptcy Costs. *The Journal of Finance*, 39(3), 629 - 642.
- Kalaycı, Ş., & Karataş, A. (2005). The relationship between common stock returns and financial ratios: A fundamental analysis in the İstanbul Stock Exchange. *Muhasebe ve Finansman Dergisi*, pp:146 - 157.
- Koh, H. C. (1992). The sensitivity of optimal cutoff points to misclassification costs of type I and type II errors in the going-concern prediction context. *Journal of Business Finance & Accounting*, 19(2), 187-197.
- Koh, H. C., & Killough, L. N. (1990). The use of multiple discriminant analysis in the assessment of the going-concern status of an audit client. *Journal of Business Finance & Accounting*, 17(2), 179-192.
- Koh, H. C., & Low, C. K. (2004). Going concern prediction using data mining techniques. *Managerial Auditing Journal*, 19(3), 462 -476.
- Kuh, E., & Meyer, J. (1955, October). Correlation and Regression Estimates when the Data are Ratios. *Econometrica*, 23(4), 400 - 416.
- Kulalı, İ. (2014, july - December). In the Light of Accounting Based Prediction Models, Comparatives of Financial Distress and Bankruptcy. *Sosyoekonomi*(2), 153-169.
- Lau, A. (1987). A Five-State Financial Distress Prediction Model. *Journal of Accounting Research*, 25(1), 127-139.

- Lev, B. (1969). Industry Averages as Targets for Financial Ratios. *Journal of Accounting Research*, Vol. 7(2 (Autumn)), pp. 290-299.
- Lev, B., & Sunder, S. (1978). Methodological issues in the use of financial ratios. *Journal of Accounting and Economics*, 1(3), 187-201.
- Lo, A. (1986). Logit versus discriminant analysis: A specification test and application to corporate bankruptcies. *Journal of Econometrics*, 31(2), 151 - 178.
- Modigliani, F., & Miller, M. H. (1958, June). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review*, 48(3), 261 - 297.
- Morrison, D. G. (1969). On the Interpretation of Discriminant Analysis. *Journal of Marketing Research*, 6(2), 156 - 163.
- Muzır, E., & Çağlar, N. (2009). The Accuracy Of Financial Distress Prediction Models In Turkey: A Comparative Investigation With Simple Model Proposals. *Anadolu University Journal of Social Sciences*, 9(2), 15-48.
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, Vol. 18(No. 1), pp. 109-131.
- Oktar, S., & Dalyancı, L. (2010). Financial Crises Theories and Financial Crises in Turkish Economy After 1990. *Marmara University Journal of the Faculty of Economic & Administrative Sciences*, 29(2), 1-22.
- Pinches, G. E., Mingo, K. A., & Caruthers, J. K. (1972). The Stability of Financial Patterns in Industrial Organizations. *The Journal of Finance*, 28(2), 389-396.
- Poston, K. M., Harmon, W. K., & Gramlich, J. D. (1994). A test of financial ratios as predictors of turnaround versus failure among financially distressed firms. *Journal of Applied Business Research*, 10(1), 41 - 56.
- Poyraz, E., & Uçma, T. (2006). The Measurement of Financial Failures of Main Exporting Sectors Against The Financial Crisis by Using Altman Z Score Model. *Muhasebe ve Finansman Dergisi*(32).
- Ruzgaresen, C. (2011). *İflas Sebepleri*. Doctoral Dissertation: Ankara University.
- Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, 74(1), 101 - 124.
- Son, H., Hyun, C., Phan, D., & Hwang, H. (2019). Data analytic approach for bankruptcy prediction. *Expert Systems with Applications*, 138, 112816.
- Stenbäck, T. (2013). Corporate default prediction with financial ratios and macroeconomic variables.
- Sun, L. (2007). A re-evaluation of auditors' opinions versus statistical models in bankruptcy prediction. *Review of Quantitative Finance and Accounting*(28), 55 - 78.
- Terzi, S. (2011). Prediction of Financial Distress Using Financial Ratios: An Empirical Research in the Food Sector. *Çukurove Üniversitesi, İİBF Dergisi*, 15(1), 1-18.

- Theil, H. (1969). On the use of Information Theory Concepts in the Analysis of Financial Statements. *Management Science*, 15(9), 459 - 480.
- Türkmen, S., & Çağıl, G. (2012). İMKB'ye Kote Bilişim Sektörü Şirketlerinin Finansal Performanslarının TOPSIS Yöntemi ile Değerlendirilmesi. *Maliye Finans Yazıları*, 26(95), pp:59 - 78.
- Uğurlu, M., & Aksoy, H. (2006). Prediction of corporate financial distress in an emerging market: the case of Turkey. *Cross Cultural Management: An International Journal*, 13(4), 277 - 295.
- Uygur, E. (2001). Krizden Krize Türkiye: 2000 Kasım ve 2001 Şubat Krizleri. 1-40.
- Van Frederikslust, R. A. (2012). *Predictability of corporate failure*. Boston: Springer Science.
- Vuran, B. (2009). Prediction of business failure: a comparison of discriminant and logistic. *Istanbul University Journal of the School of Business Administration*, 38(1), 47-65.
- Warner, J. B. (1977). Bankruptcy Costs: Some Evidence. *The Journal of Finance*, 32(2), 337 - 347.
- White, M. J. (1983). Bankruptcy Costs and the New Bankruptcy Code. *The Journal of Finance*, 38(2), 477 - 488.
- Whittington, G. (1980). Some basic properties of accounting ratios. *Journal of Business Finance & Accounting*, 7(2), 219 - 232.
- Wilcox, J. W. (1971). A Simple Theory of Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 9(2), 389 - 395.
- Yurdakul, M., & İç, Y. T. (2003). Türk Otomotiv Firmalarının Performans Ölçümü ve Analizine Yönelik TOPSIS Yöntemini Kullanan Bir Örnek Çalışma. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 18(1), 1-18.
- Zavgren, C. (1985). Assessing the vulnerability to failure of American industrial firms: a logistic analysis. *Journal of Business Finance & Accounting*, 12(1), 19-45.
- Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*, 22, 59 - 82.